

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2004, 07:07:47 ; Search time 1236.96 Seconds
(without alignments)

6326.134 Million cell updates/sec

Title: US-09-977-053-6

Perfect score: 10219
Sequence: 1 MWPRFLAFCCWGLALVSGWAT.....GVTKITCLSGEWNHLPYC 1842

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2 1/USPTO spool_p/US0977053/runat 06052004 075942 18178/app query.fasta 1.5710
-DB=N Geneseq 29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=100 -LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=25
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0977053 @CNG 1 1 4042 @runat 06052004 075942 18178 -NCFU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002s:.*
7: geneseqn2003as:.*
8: geneseqn2003bs:.*
9: geneseqn2003cs:.*
10: geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10214	100.0	11546	8 AAD57253	Aad57253 Human CGD
2	10208	99.9	10878	6 AAD32025	Aad32025 Human C3b
3	9862	96.5	11152	6 ABT08491	Abt08491 Human nov
4	9812	96.0	11158	6 ABT08490	Abt08490 Human nov
5	8652.5	84.7	11230	6 AAD32026	Aad32026 Mouse C3b
6	7537	73.8	5124	4 AAH16667	Aah16667 Human C3b
7	5485	53.7	3448	6 ABZ11152	Abz11152 Human pol
8	4730	46.3	3128	6 ABK54124	Abk54124 CDNA enco

C	9	3133	30.7	2064	6	ABK35664	ABK35664	CDNA sequ
	10	3073	30.1	1969	4	AAH16567	AAH16567	Human cDN
	11	2479	24.3	1696	6	ABK54146	ABK54146	CDNA enco
	12	2374	23.2	1760	3	AZ294659	AZ294659	Human EGF
	13	1977	19.3	1408	6	ABK35663	ABK35663	CDNA sequ
	14	1498	14.7	1892	4	AAS26857	AAS26857	Human cDN
	15	1447	14.2	765	4	AAH08402	AAH08402	Human cDN
	16	1354	13.2	10489	4	ABL14889	ABL14889	Drosophil
	17	1341	13.1	706	7	ACA57318	ACA57318	Human adi
C	18	1328	13.0	13484	4	ABL14888	ABL14888	Drosophil
	19	1304	12.8	801	4	AAH07752	AAH07752	Human cDN
	20	1278	12.5	680	4	AAS28962	AAS28962	CDNA enco
	21	1278	12.5	680	4	AAS26930	AAS26930	Human cDN
	22	1278	12.5	680	4	ABA06507	ABA06507	Human cDN
	23	1278	12.5	680	4	AAS31580	AAS31580	CDNA enco
	24	1278	12.5	680	6	ABT07797	ABT07797	Novel hum
	25	1278	12.5	680	6	ABV83844	ABV83844	Human pol
	26	1237	12.1	3262	6	ABS51807	ABS51807	Human mdd
	27	998	9.8	598	2	AAV87221	AAV87221	EST clone
C	28	983	9.6	675	5	AAD02459	AAD02459	Mouse P10
	29	975	9.5	2697	7	ABX34482	ABX34482	Human mdd
	30	914.5	8.9	6728	6	AAD28942	AAD28942	Human MOL
	31	914.5	8.9	6728	9	ADD18193	ADD18193	Human mol
	32	914.5	8.9	9723	2	AA155430	AA155430	Human Not
	33	914	8.9	8113	4	ABL06899	ABL06899	Drosophil
	34	893.5	8.7	7410	6	AAD28941	AAD28941	Human MOL
	35	893.5	8.7	7410	9	ADD18191	ADD18191	Human mol
	36	872	8.5	7596	7	AAD56128	AAD56128	Mouse Not
	37	872	8.5	7596	8	ADA02491	ADA02491	Mouse Not
	38	872	8.5	7596	9	ADB72229	ADB72229	Mouse Not
	39	872	8.5	8064	7	AAD56127	AAD56127	Mouse Not
	40	872	8.5	8064	8	ADA02490	ADA02490	Mouse Not
	41	872	8.5	8064	9	ADB72228	ADB72228	Mouse Not
	42	869.5	8.5	8221	7	ABT42465	ABT42465	Toxicity
	43	860	8.4	7673	7	ABS55875	ABS55875	Human pol
	44	853	8.3	7332	6	ABL40768	ABL40768	Constitut
C	45	847	8.3	37322	6	ABL06898	ABL06898	Drosophil
	46	838	8.2	763	6	ABS70453	ABS70453	Human bon
	47	837.5	8.2	2460	6	AAD28494	AAD28494	Human ext
	48	831.5	8.1	7615	6	AAD56130	AAD56130	Human NOT
	49	831.5	8.1	7615	8	ADA02493	ADA02493	Human NOT
	50	831.5	8.1	7615	9	ADB72231	ADB72231	Human NOT
	51	829	8.1	8091	2	AAV57001	AAV57001	Human Not
	52	829	8.1	8091	6	ABZ34974	ABZ34974	Human gen
	53	829	8.1	8257	6	ABS70408	ABS70408	Human bon
	54	764.5	7.5	6677	6	AAH00989	AAH00989	Mus muscu
	55	764.5	7.5	6677	6	AAI72830	AAI72830	Mouse Not
	56	762	7.3	5617	2	AAV57163	AAV57163	Partial h
	57	745	7.3	3974	7	ABX75297	ABX75297	Mouse not
	58	745	7.3	3974	7	AAD52524	AAD52524	Murine no
	59	745	7.3	3974	7	ABZ81794	ABZ81794	Mouse not
	60	745	7.3	4855	6	ABK35564	ABK35564	Gene JAG1
	61	745	7.3	6464	2	AAT40090	AAT40090	Human Ser
	62	745	7.3	6464	3	AZ49096	AZ49096	Human ser
	63	744	7.3	4208	2	AAV15201	AAV15201	Human ser
	64	744	7.3	5590	2	AAV63753	AAV63753	Human JAG
	65	744	7.3	5590	8	ACD06196	ACD06196	Human cDN
	66	744	7.3	5896	7	ABZ34802	ABZ34802	Coding se
	67	744	7.3	5929	7	ACA04012	ACA04012	CDNA down
	68	743	7.3	4208	2	AAT70175	AAT70175	Prolifera
	69	739.5	7.2	4146	9	AAD60679	AAD60679	Human CRU
	70	739.5	7.2	4361	9	AAD60677	AAD60677	Human CRU
	71	739.5	7.2	4413	7	ACD05903	ACD05903	Novel hum
	72	739.5	7.2	4950	9	AAD60678	AAD60678	Human CRU
	73	739	7.2	5942	9	ADD14711	ADD14711	Human 8rc
	74	737	7.2	5458	2	AAV03674	AAV03674	Human Jag
	75	733	7.2	4200	6	ABV99395	ABV99395	Human NOV
	76	727	7.1	3825	9	AAD60686	AAD60686	Human CRU
	77	727	7.1	3921	9	AAD60683	AAD60683	Human CRU
	78	721	7.0	4200	6	ABV99396	ABV99396	Human NOV
	79	716.5	7.0	3786	9	AAD60684	AAD60684	Human CRU
	80	714.5	7.0	2757	4	AAS26875	AAS26875	Human cDN
C	81	713	7.0	462	4	AAS26893	AAS26893	Human cDN

c	82	713	7.0	462	4	AAS26819	Human cDN
	83	712	7.0	7216	4	ABL11001	Drosophila
	84	708	6.9	5878	7	ABX75298	Human not
	85	708	6.9	5878	7	AAD52525	Human not
	86	708	6.9	5878	7	ABZ81795	Mouse not
	87	691.5	6.8	5848	9	AAD62276	Rat IRDBP
	88	690	6.8	3579	2	AAT40092	Chick Ser
	89	690	6.8	3582	3	AAT40098	Chick Ser
	90	688	6.7	2575	6	ABQ72635	Human MDD
	91	687.5	6.7	4926	4	AAF30860	Rat insul
	92	687.5	6.7	4926	9	AAD62243	Rat IRDBP
	93	686.5	6.7	1717	6	ABK71805	Human dit
	94	686.5	6.7	3612	9	AAD60685	Human CRU
	95	684	6.7	448	6	ABN15878	Human ORF
	96	679.5	6.6	4049	7	AAD57530	Murine no
	97	679.5	6.6	4050	7	ABX75300	Mouse not
	98	679.5	6.6	4050	7	ABZ81797	Human not
	99	673	6.6	4333	8	ACD06197	Human cDN
	100	672	6.6	4749	9	ADD00948	Human Jag

ALIGNMENTS

RESULT 1
AAS26819 standard; cDNA; 11546 BP.

XX	AC	AAS26819	
XX	DT	06-NOV-2003	(first entry)
XX	DE	Human CGDD-33	cDNA.
XX	KW	Human; cell growth, differentiation and death protein; CGDD; leukaemia;	
XX	KW	neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;	
XX	KW	muscular disorder; myotonic dystrophy; catatonias; endocrine disorder;	
XX	KW	diabetes; Grave's disease; cancer; immunological disorder; scleroderma;	
XX	KW	systemic lupus erythematosus; allergy; Crohn's disease; renal disorder;	
XX	KW	gastrointestinal disorder; Goodpasture's syndrome; infection; cirrhosis;	
XX	KW	cardiovascular disorder; atherosclerosis; hepatic disease; transgenic;	
XX	KW	transgenic animal; gene therapy; neuroprotective; relaxant; cytosolic;	
XX	KW	dermatological; immunosuppressive; cerebroprotective; anticonvulsant;	
XX	KW	antibacterial; antiparasitic; fungicide; virucide; uropathic; cardiant;	
XX	KW	protozoacide; nootropic; gene; ss.	
XX	OS	Homo sapiens.	
XX	XX	Key	Location/Qualifiers
XX	FT	CDS	298..11010
XX	FT		/*tag= a
XX	FT		/product= "Human CGDD-33 protein"
XX	PN	WO2003050253-A2.	
XX	PD	19-JUN-2003.	
XX	XX	04-DEC-2002; 2002WO-US039133.	
XX	PR	07-DEC-2001; 2001US-0340747P.	
XX	PR	20-DEC-2001; 2001US-0342761P.	
XX	PR	15-JAN-2002; 2002US-0349705P.	
XX	PR	06-FEB-2002; 2002US-0354764P.	
XX	PR	12-FEB-2002; 2002US-0356216P.	
XX	XX	(INCY-) INCYTE GENOMICS INC.	
XX	XX	Griffin JA, Ramkumar J, Emerling BM, Kable AB, Elliott VS;	
XX	PI	Marquis JP, Baughn MR, Yue H, Lee EA, Becha SD, Tang YT;	
XX	PI	Tran UK, Swarnakar A, Lee S, Ison CH, Hafalia AJA, Tran B;	
XX	PI	Sprague WW, Lee SY, Khare R, Gandhi AR, Gietzen KJ, Bhatia U;	
XX	PI	Burrill JD, Blake JJ, Ho A, Zheng W;	
XX	DR	WPI; 2003-532903/50.	

DR	P-PSDB; AAD37944.
XX	New CGDD polypeptides, useful for diagnosing, preventing, and treating
PT	disorders associated with an abnormal expression or activity of CGDD,
PT	e.g. neuromuscular, immunological, cardiovascular disorders, cancer
PT	and/or infections.
XX	Claim 5; Page 296-299; 299pp; English.
PS	The present invention relates to novel cell growth, differentiation and
XX	death (CGDD) proteins and polynucleotides encoding them. The sequences of
CC	the invention are useful in diagnosing, preventing and treating disorders
CC	associated with an abnormal expression or activity of CGDD such as
CC	neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's
CC	disease), muscular disorders (e.g. myotonic dystrophy, catatonias),
CC	endocrine disorders (e.g. diabetes, Grave's disease), cancers (e.g.
CC	leukemia, cervical or breast cancers), immunological disorders (e.g.
CC	scleroderma, systemic lupus erythematosus, allergies), gastrointestinal
CC	disorders (e.g. Crohn's disease), renal disorders (e.g. Goodpasture's
CC	syndrome), infections (e.g. viral, bacterial, fungal, parasitic,
CC	protozoal, helminthic), cardiovascular disorders (e.g. atherosclerosis)
CC	and hepatic diseases (e.g. cirrhosis). The polynucleotides can be used to
CC	create humanised animals or transgenic animals to model human diseases.
CC	The invention is also used in gene therapy. The present sequence is human
CC	CGDD-33 cDNA
XX	SQ Sequence 11546 BP; 3008 A; 2726 C; 2836 G; 2976 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	0 Length: 11546
Score:	10214.00 Matches: 1840
Percent Similarity:	100.00% Conservative: 2
Best Local Similarity:	99.99% Mismatches: 0
Query Match:	99.99% Indels: 0
DB:	8 Gaps: 0
US-09-977-053-6 (1-1842) x AAD57253 (1-11546)	
Qy	1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
Db	298 ATGTGGCCCTCGCTGGCTTTTGTGTGGGGTCTGGGGCTGGTTTCGGGGTGGGGGACC 357
Qy	21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
Db	358 TTTTACAGACAGATGTCTCCCGTGGCGCAATTTTACGCTTCCGCTCTTCCCGAGACCGGGCCC 417
Qy	41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
Db	418 GGGGCCCCCGGAGTATCTCCCGCGCCGCTCTCTGGCGAGCGAGCGCGGAGCAGA 477
Qy	61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
Db	478 GTGGAGCGGCTGGGCGAGGCGTTCGGCGAGCGGCTGGCTGGCGAGGCTCAGCGAG 537
Qy	81 ArgLeuGluLeuValPheLeuValAspAspSerSerSerValGlyGluValAsnPheArg 100
Db	538 CGCTTGGAGCTTGTCTTCTCTGTGTGTATTTCTTCCAGCTGGGCGAGTCAACTTCGCG 597
Qy	101 SerGluLeuMetPheValLargLysLeuLeuSerAspPheProValValProThrAlaThr 120
Db	598 AGCGAGCTCATGTTCTGTTCGCGAAGCTGTCTGTCCGACTTCCCGTGGTGGCCACGCGCAG 657
Qy	121 ArgValAlaIleValThrPheSerSerSerLysAsnTyrValValProArgValAspTyrIle 140
Db	658 CGCGTGGCCATCTGTGACCTTCTCTGTCAGAACTACGTGGTGGCGCGCTCGATTACATC 717
Qy	141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAlaIle 160
Db	718 TCCACCCCGCGCGCGCGCAGCAGCAGTGGCGCTGTCTCTCCAGAGAGATCTCTGCCATC 777
Qy	161 SerTyrArgGlyGlyGlyThrThrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeu 180
Db	778 TCTTACCGAGGTGGCGGCACTTACACCAAGGGCGGCTTCCAGCAAGCGCGCAAAATCTTT 837

181 LeuHisAlaArgGluAsnSerThrIleValValPheLeuIleThrAspGlyTyrSerAsn 200
182 LeuHisAlaArgGluAsnSerThrIleValValPheLeuIleThrAspGlyTyrSerAsn 200
838 CTTCAATGCTAGAGAAACCTCAAAAGATGTATTTCTCATCTGATGGATATTCAT 897
201 GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr 220
202 GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr 220
898 GGGGAGAGCCCTAGACCAATTCAGCGTCACTGCGAGATTCAGGAGTGGAGATCTTCACT 957
221 PheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIleGlu 240
222 PheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIleGlu 240
958 TTTGGCATATGGCAAGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 1017
241 GluHisCysValTyrLeuLeuHisSerPheGluGluPheGluAlaAlaArgAlaLeu 260
242 GluHisCysValTyrLeuLeuHisSerPheGluGluPheGluAlaAlaArgAlaLeu 260
1018 GAGCACTGTTTACCTCTACACAGTTTTTGAAGAAATTTGAGGCTTTAGCTCGCGGCAATTG 1077
261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
262 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
1078 CATGAAGATCTACCTTCTGGGAGTTTTTATTCAAGATGATATGGTCCACTGCTCATATCTT 1137
281 CysAspGluGlyAspCysAspArgMetGlySerCysIleGlyThrHisThr 300
282 CysAspGluGlyAspCysAspArgMetGlySerCysIleGlyThrHisThr 300
1138 TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGGAAGCTGCMAATGTGGGACACACACA 1197
301 GlyHisPheGluCysIleCysGluGlyGlyTyrTyrGlyGlyLeuGlnTyrGluCys 320
302 GlyHisPheGluCysIleCysGluGlyGlyTyrTyrGlyGlyLeuGlnTyrGluCys 320
1198 GGCATTTTGTAGTGATCTGTGAAAGGGGTATTACGGAAAGGCTGCAGATGATGATGTC 1257
321 ThrAlaCysProSerGlyThrTyrIleGlyProGluGlySerProGlyIleSerSerCys 340
322 ThrAlaCysProSerGlyThrTyrIleGlyProGluGlySerProGlyIleSerSerCys 340
1258 ACAGCTTGCCCATCGGGGACATACAAACCTGGAAGGCTCACAGGAGGATCAGCAGTTGC 1317
341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
342 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
1318 ATTCCATGCTCTGATGAATAATCACACCTCTCCACCTGGGAAGCACATCCCTCGAAGACTGT 1377
361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
362 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
1378 GTCTGAGAGAGGATACAGGGCATCTGGCCAGACTGTGGAATCTGTCCACTGCCCTGCC 1437
381 LeuIleProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
382 LeuIleProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
1438 CTGAAGCTTCCGAAATGGTTACTTTATCCAAACACCTTGCAACACCACTTCATGCA 1497
401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
402 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
1498 GCTGTGGGGTCCGATGTCCACCTGATTTGATCTTTGTGGGAAGCAGCATCATCTTATGT 1557
421 LeuProAsnGlyLeuThrSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
422 LeuProAsnGlyLeuThrSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
1558 CTACCAATGGTTGTGGTCCGGTTTCAAGAGCTACTCGAGATGAAGACATGCTCTCAT 1617
441 LeuArgGlnProIleHisGlyHisIleSerCysSerThrArgGluMetLeuTyrIleThr 460
442 LeuArgGlnProIleHisGlyHisIleSerCysSerThrArgGluMetLeuTyrIleThr 460
1618 CTCCGCGAGCGAACAATGGCCACATCAGCTGTTCTCAAGGGGAATGTTATATAAGACA 1677
461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspIleLeuThrCys 480
462 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspIleLeuThrCys 480
1678 ACATGTTTGGTGTGCTGATGAAGGTACAGACTAGAGGCGAGTGAATAGCTTACTTGT 1737
481 GlnGlyAsnSerGlnThrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
482 GlnGlyAsnSerGlnThrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
1738 CAAGAAACAGCCATGTGGGATGGGCGAAGCCCGGTGTGGAGCGCCACTGTTCCACC 1797
501 PheGlnMetProIleAspValIleIleSerProHisAsnCysGlyIleGlnProAlaIle 520
502 PheGlnMetProIleAspValIleIleSerProHisAsnCysGlyIleGlnProAlaIle 520
1798 TTTTCAGATGCCAAGATGTATCATATCCCTCCACACTGTGGCAAGCAGCAGCCAAA 1857
521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValIleGlu 540
522 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValIleGlu 540
1858 TTTGGGAGCATCTGCTATGTAAGTTGGCCGCAAGGGTTCAATTTTATCTGGAGTCAAGAA 1917

541 MetLeuArgCysThrThrSerGlyIleThrAsnValGlyValGlnAlaAlaValCysIle 560
542 MetLeuArgCysThrThrSerGlyIleThrAsnValGlyValGlnAlaAlaValCysIle 560
1918 ATGCTGAGATGTACCACTTCTGGAAATGGAATGTGGAGTTCAGGCGAGCTGTGTGTA 1977
561 AspValGluAlaProGlnIleAsnCysProIleAspIleGluAlaIleThrLeuGluGln 580
562 AspValGluAlaProGlnIleAsnCysProIleAspIleGluAlaIleThrLeuGluGln 580
1978 GAGCTGGAGGCTCTCTCAATCACTGTCTTAAGACATAGAGGCTAAGACTCTGGAACAG 2037
581 GlnAspSerAlaAsnValThrThrGlnIleProThrAlaIleAspAsnSerGlyGluIle 600
582 GlnAspSerAlaAsnValThrThrGlnIleProThrAlaIleAspAsnSerGlyGluIle 600
2038 CAGATCTTGGCCCAATGTATCTGGCAGATTCACACAGCTAAGACAACTCTGGTGA 2097
601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
602 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
2098 GTGTCACTCCAGCTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGT 2157
621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
622 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
2158 GCTATCTGATACAGGCAACTGACCTATTCGGCAACACAGCCAGCTGCAATTTTCCATATC 2217
641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
642 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
2218 AAGTTTATGATGACAGAACCCCTGTATAGACTGGTGAGATCTCCACCTCCCGTCCAG 2277
661 ValSerGluIleValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
662 ValSerGluIleValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
2278 GTCTCGAGAGGTACATGCGCGAAGCTGGGATGAGCTCAGTTCTCAGACAACTCAGG 2337
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyThr 700
682 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyThr 700
2338 GCTGAATGCTCATTAACAGAGTCAATACAGAGGAGACCTTTTCCCTCAAGGGGAGACT 2397
701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
702 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
2398 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATTAACAGGACATGTATATCATATT 2457
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
722 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
2458 GTCATANAAAGTTCTCCCTGTGAATTCATTCACACTGTAAATGGGATTTATATGTC 2517
741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
742 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
2518 ACTCCAGATAATATCTGGAGTCAACTGTATTAATCTTGTGGAGGCTATGATTTTCA 2577
761 GluGlySerThrAspIleTyrCysAlaTyrGluAspGlyValTyrIleProThrTyr 780
762 GluGlySerThrAspIleTyrCysAlaTyrGluAspGlyValTyrIleProThrTyr 780
2578 GAAGGCTCTACTGACAGATTAATTTGTCTTATGAAGATGGCGTCTGGAACCAACATAT 2637
781 ThrThrGluTrpProAspCysAlaIleLysArgPheAlaAsnHisGlyPheLysSerPhe 800
782 ThrThrGluTrpProAspCysAlaIleLysArgPheAlaAsnHisGlyPheLysSerPhe 800
2638 ACCCTGATGGCCAGACTGTGCCAANAACGTTTTTGCACCAACCAAGTTCAGTCTTT 2697
801 GluMetPheTyrIleAlaAlaArgCysAspAspThrAspLeuMetIleLysPheSerGlu 820
802 GluMetPheTyrIleAlaAlaArgCysAspAspThrAspLeuMetIleLysPheSerGlu 820
2698 GAGATGTTCTCAAAAGCAGCTGTTGTGATGACACAGATCTGATGAAGAGTTTCTGAA 2757
821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
822 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
2758 GCATTTGACAGCAGCTGCGGAAAATGGTCCCATCATTTTGTAGTATGATGAGAGGACAT 2817
841 AspCysArgLeuGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
842 AspCysArgLeuGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
2818 GACTGCAGACTGGAGGAGAACCTGACCAANAATATTGCTAGATATATAATTATGACTAT 2877
861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
862 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
2878 GAAATGCTTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 2937
881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
882 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
2938 TACGATGACTTCTGAGCACTGTGCAAGAAACAGCCACAGCATCGGCAATGCCAAGTCC 2997
901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920

Db	2998	TCACGGATTAAGAGAGTGCCTTATCTGACATATAAATAAGTTAAATTTTAAATC	3057	Db	4078	TACACAGTGCAGCGGTGTGAAGAAAATATAAATGAGTGTAGTCCAGTCTCTTGTTTAAAT	4137
Qy	921	ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluIrrpGluAsnGln	940	Qy	1281	LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal	1300
Db	3058	ACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAA	3117	Db	4138	AAAGGAATCTGTGTGATGGTGTGCTATCGTATCGTATCGTATCGTATCGTATCGT	4197
Qy	941	ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLysLysArgThrLeuAsnLysAsp	960	Qy	1301	GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla	1320
Db	3118	CGACTCTCTCAGACATTTGGAACTATCACAATAAATGAAAGGACTCTCAACAAAGAC	3177	Db	4198	GGCTGCAATTTGAAACAGAGAGTCAATGATGCCAGTCAACCCATGCTTAATAATGCA	4257
Qy	961	ProMetTyrSerPheGlnLeuAlaSerGluIleLeuLeuAlaAspSerAsnSerLeuGlu	980	Qy	1321	ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr	1340
Db	3178	CCCATGTATCTCTTACCTTGCATCAGAAATCTTATAGCCGACAGCAATTCATTAGNA	3237	Db	4258	GTCTGTGAAGACCAAGTTGGGGGATTCTTGTGCAAAATGCCACCTGGATTTTGGGTACC	4317
Qy	981	ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal	1000	Qy	1341	ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys	1360
Db	3238	ACAAAAGAGCTTCCCTCTCTGCAGACACAGCTCAGTGTCTGAGAGGGCGTATGTGTCT	3297	Db	4318	CGATGTGAAAGAGAGCTCGATGAGTGTCTCAGTGCAGCATGCAAAATGGAGCTACCTGT	4377
Qy	1001	AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle	1020	Qy	1361	LysAspGlyAlaAsnSerPheArgCysAlaAlaGlyPheThrGlySerHisCys	1380
Db	3298	AAITGCCCTTTGGGAACCTATTATAATCTGGAACATTTTACCTGTGAAAGCTGCCGGATC	3357	Db	4378	AAAGAGGTGCATAGCTTCCAGATCCCTGTGTGAGTGGCTTCCAGGATCACACTGT	4437
Qy	1021	GlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr	1040	Qy	1381	GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp	1400
Db	3358	GGATCTATCAAGATGAAGAGGGCACTTGAGTGCAGAGCTTTGCCCTCTGGGATGTAC	3417	Db	4438	GAATTTGAACATCAATGAATGTCACTTAATCCATGTAGAAATCAGGCCACCTGTGTGGAT	4497
Qy	1041	ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr	1060	Qy	1401	GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr	1420
Db	3418	ACGGAATATATCCATTCAAGAAACATCTCTGAATGTAAAGCTCAGTGTAAACAGGCACC	3477	Db	4498	GAATTTAAATTCATACAGTTGTAAATGTCAAGCAGATTTTTCAGGCAGAAAGGTGTGAAACA	4557
Qy	1061	TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys	1080	Qy	1421	GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet	1440
Db	3478	TACTCATACAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCACCTTATCAGCCAAA	3537	Db	4558	GAACAGTCTACAGGCTTTAACTGTGATTTTGAAGTTTCTGGCATCTATGGATATGTCATG	4617
Qy	1081	PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla	1100	Qy	1441	LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSerSer	1460
Db	3538	TTTGTTTCCCGAGCTGCTCTGTGTCCAGAAACACCTCACTGTGTGAAAGAGGAGGCC	3597	Db	4618	CTAGATGGCATGCTCCCATCTCTCACTGTAACTGTACCTTCTGGATGAAATCCTCT	4677
Qy	1101	ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu	1120	Qy	1461	AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr	1480
Db	3598	GTGAACATTTCTGATGTGGAGTCTCTTGTCCAGAGGAAATTTCTCGCGTCTCGGGTTA	3657	Db	4678	GACGACATGAACTATGGAACACCAATCTCTATGAGTTGATAGCGGAGGACAAATACC	4737
Qy	1121	MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys	1140	Qy	1481	LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGlyLysIleThr	1500
Db	3658	ATGCCCTGTACCCATGTCCTGTGACTATTACCAACCTTAATGAGGAGGAGCCCTTCGC	3717	Db	4738	TTGCTCTGACTGATTAACGGCTGGGTCTTATGTGATGATGCGGAGGAGGATAACA	4797
Qy	1141	LeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGluCys	1160	Qy	1501	AsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrThrThrSerAla	1520
Db	3718	CTGGCCCTGTCTTTATGGAACTACCCCATCTGCTGTGTTCCAGATCGCTCACGAATGT	3777	Db	4798	AACTGTCTCTCGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGCAAGTGCC	4857
Qy	1161	SerSerPheSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeu	1180	Qy	1521	AsnGlyIleThrLysValTyrIleAspGlyLysLeuSerAspGlyValAlaGlyLeuSer	1540
Db	3778	TCAAGTTTATGTTCAACTTCTCAGCGCAGAGGAAAGTGTGTGCGCCCTCTCTCT	3837	Db	4858	AATGGCATCTGGAAGTCTATATCGAATGGGAAATATCTGACGGTGTGCTGGCCCTCT	4917
Qy	1181	GlyHisIleLysArgHisGluIleSerSerGlnValPheHisGluCysPheAsn	1200	Qy	1541	ValGlyLeuProIleProGlyGlyValAlaLeuValLeuGlyGlnGlnAspLysLys	1560
Db	3838	GGACATATAAAGAGCGCATGAATCAGCAGTCAAGTGTGTTTCCATGAATGCTCTTAAAC	3897	Db	4918	GTGTTTGGCCATACCTGT	4977
Qy	1201	ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro	1220	Qy	1561	GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuThr	1580
Db	3898	CTTTGCCACATAGTGGAACTTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3957	Db	4978	GGAGAGGATTCAGCCAGCTGAGTCTTGTGTGGCTCCATAGGCCAGCTCAACCTCTGG	5037
Qy	1221	LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys	1240	Qy	1581	AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu	1600
Db	3958	CTTGGATATACAGGCTTAAAGTGTGAAACAGACATCGATGAGTGCAGCCCATCTCTTGC	4017	Db	5038	GACTATGTCTGTCTCCACAGCGGTGAGTCACTGGCTTACCTCTCTGCCAGAGGAACTC	5097
Qy	1241	LeuAsnGlnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly	1260	Qy	1601	SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys	1620
Db	4018	CTCAACAATGAGTGTGTGAAAGCTATGTGGGAAATTCATTGTGAGTGCCTATCAGGT	4077	Db	5098	AGTAAAGGAACGTGTAGCATGGCTGATTTCTTGTGAGGAATTTGGGGAAGTGAAG	5157
Qy	1261	TyrThrGlyGlnArgCysGluAlaAsnIleAsnGluCysSerSerProCysLeuAsn	1280	Qy	1621	IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis	1640
				Db	5158	ATCGATTTCTAGAGCATATTTTGTCTGATTTGCCACCGCTTAGGAGGGTCACTGCTCAT	5217

Db 371 |||||CGCGTGGCCATCGTGACCTTCTGTCGAAGAACTACGTGGTGGCCGCGTCAATTACATC 430
Qy 141 SerThrArgArgAlaAargGlnHisLysCysAlaLeuLeuLeuGlnGluLeuProAlaIle 160
Db 431 TCCACCGCGCGCGCGCCAGCACCAAGTGGCGCTGCTCTCCCAAGAGATCCCTGCCATC 490
Qy 161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeu 180
Db 491 TCCTACCGAGGTGGCGGCACCTACACCAAGGCGCTTCCAGCAAGCGCGCAATTCTT 550
Qy 181 LeuHisAlaArgGluAsnSerThrLysValPheLeuLeuLeuThrAspGlyTyrSerAsn 200
Db 551 CTTCAATGCTAGAGAAACTCAACAAAGTTGATTTCTCATCACTGATGATATTCAT 610
Qy 201 GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr 220
Db 611 GGGGAGAGCCCTAGACCAATTCAGCGCTCACTGCGAGATTTCAGGAGTGGAGATCTTCACT 670
Qy 221 PheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
Db 671 TTTGGCATATGGCAAGGAACTTCGAGAGCTGAATGATGCTTCCACCCCAAGGAG 730
Qy 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeu 260
Db 731 GAGCACTGTTACTCTCTACACAGTTTGAAGATTTGAGGCTTTAGCTCGCGGCGATTG 790
Qy 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
Db 791 CATGAAGATCTACCTTCTGGGAGTTTATTTCAAGATGATATGGTCCACTGCTCTTATCTT 850
Qy 281 CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
Db 851 TGTGATGAAGCGAAGGACTGCTGTGACCGAATGGGAAGCTGCAAAATGGGGACACACACA 910
Qy 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
Db 911 GGCCATTTTGTAGTGCACTGTGTGAAAAGGGTATTACGGGAAGGTTCTGCAGTATGATGC 970
Qy 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
Db 971 ACAGCTTCCCATCGGGACATACAAACCTGAAGGCTCACAGGAGGAATCAGCAGTTGC 1030
Qy 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
Db 1031 ATTTCATGTCTCTGATGAAAAATCACACCTCTCCACCTGGGAAGCACATCCCTCAAGACTGT 1090
Qy 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
Db 1091 GTCTGCAGAGAGGGATACGGGCATCTGGCCAGACTGTGAACCTGTCACTGCCCTGCC 1150
Qy 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
Db 1151 CTGAAGCCCTCCGAAAAATGGTTACTTTATCCAAAAACCTTGCAACAAACCTTCAATGCA 1210
Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
Db 1211 GCCTGTGGGTCCGATGTCAACCTGATTTGATCTTGTGGGAAGCAGCATCATCTATGT 1270
Qy 421 LeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
Db 1271 CTACCCATGTTTGTGGTCCGGTTCCAGAGACTACTGCAGAGTAAGAATGCTCTCAT 1330
Qy 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
Db 1331 CTCCGCCAGCCGAAACATCGCCACATCAGCTGTCTTCAAGGGAATGTTATATAGACA 1390
Qy 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
Db 1391 ACATGTTTGGTTCCTGTGATGAAGGGTACAGCTAGAGAGGAGTATGATAGCTTACTTGT 1450
Qy 481 GlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
|||

Db 1451 CAAGGAAACAGCCAGTGGGATGGCCAGAAACCCCGGTGTGTGGAGCGCCACTGTGTTCCACC 1510
Qy 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
Db 1511 TTTTCAGATGCCCAAGATATCAT 1570
Qy 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
Db 1571 TTTGGAGAGATCTGCTATGTAAGTTGGCCCAAGGGTTCAATTTATCTGGAGTCAAGAA 1630
Qy 541 MetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAlaValCysLys 560
Db 1631 ATGCTGAGATGTACCACTTCTGGAATAATGGAATGTCGAGTTTCAGGAGCTGTGTGTAA 1690
Qy 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
Db 1691 GACGTGGAGGCTCTCTCAAAATCAACTGCTCTAAGGACATAGAGGCTTAAGACTCTGGAA 1750
Qy 581 GlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyLysLys 600
Db 1751 CAAAGATTCTGCCAATGTTTACCTGGCAGATTCCACAGCTTAAGACAACCTCTGGTGAAG 1810
Qy 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
Db 1811 GTGTCACTCCAGCTTCATCCAGCTTTCACCCCACTTACCTTTTCCCACTGGAGATGT 1870
Qy 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
Db 1871 GCTATCGTATACACGGCAACTGACCTATCCGGCAACACGAGCCAGCTGCAATTTTCCATATC 1930
Qy 641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
Db 1931 AAGGTTATGTATGACAGAACCACTTGTCTATAGACTGGTGCAGATCTCCACCTCCGTCAG 1990
Qy 661 ValSerGluLysValHisAlaIleSerThrAspGluProGlnPheSerAspAsnSerGly 680
Db 1991 GTCTCGGAGAGGTACATGCGCAAGCTGGGATGAGCCTCAGTTCTCAGACAACCTCAGGG 2050
Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
Db 2051 GCTGAATTTGCTCATTTACCAAGAGTCAATACACAGGAGACTTTTCCCTCAAGGGAGACT 2110
Qy 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Db 2111 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATAT 2170
Qy 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
Db 2171 GTCATAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATATGC 2230
Qy 741 ThrProAspAsnThrGlyValAlaAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
Db 2231 ACTCAGATAATACTGGAGTCACTGTATTAACCTTGTGGAGGGCTATGATTTTCACA 2290
Qy 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
Db 2291 GAAGGCTCTACTGACAGATATTATTGTCTTATGAAGATGGCGCTCTGGAACCAACATAT 2350
Qy 781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Db 2351 ACCACTGATGGCAGACTGTGCCAAAAACGTTTTGTCTAACCCAGGGGTTCAAGTCTCTT 2410
Qy 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
Db 2411 GAGATGTTCTACAAAGCACTCGTTGTGATGACACAGACTGTGATGAAGATTTTCTGAA 2470
Qy 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
Db 2471 GCATTTGACAGACCCCTGGGAAAAATGGTCCCATCATTTTGTAGTGAAGAGATTTCTGAA 2530
Qy 841 AspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
Db 2531 GACTGCAGACTGGAGGAGAACCTGACCAAAAAATATTTGCTTGAATATATATATGACTAT 2590

QY 861 GluAenGlyPheAlaIleGlyProGlyGlyTTPGlyAlaAlaAenArgLeuAspTyrSer 880
DB 2591 GAAATGGCTTTGGCAATTGGACACAGGTGGCTGGGGTGCAGCTAAATAGAGCTGGATTACTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGlnThrAlaThrSerIleGlyAsnAlaIysSer 900
DB 2651 TACGATGACTTCTGGACACTGTGCAGAGAAACAGCCACAGCATCGGCATGCCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
DB 2711 TCACGGATTAAAGAGAGTGCCTCATTTATCTGACTATAAAATTAAGTTAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTTPGluAenGlnGln 940
DB 2771 ACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAATACAGAA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
DB 2831 CGACTCTTTCAGACATTTGGAACTATCACAAATTAACCTGAAAGGACTTCAACAAGAC 2890
QY 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
DB 2891 CCCATGATTCTTTCAGCTTGCATCAGAAATACTTTATAGCCGACAGCAATTCATTAGAA 2950
QY 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
DB 2951 ACAAAAAGGCTTCCCTCTTCGACAGCAGGCTCAGTGTGAGAGGGCTATGTGTGTC 3010
QY 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
DB 3011 AATGTCCCTTTGGGAACCTTATTATATCTGGAACATTTTACCTGTGAAGCTGCCGATC 3070
QY 1021 GlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
DB 3071 GCATCTCATAGATGAAGAGGCAACTTGTAGTCAAGCTTTGCCCTCTGGGATGTAC 3130
QY 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
DB 3131 ACGGAATATATTCATTCAAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAGGCACC 3190
QY 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
DB 3191 TACTCATACGTGGACTTGAGCTTGTGAATCTGTCTCCACTGGGCATTTATCAGCCAAA 3250
QY 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
DB 3251 TTTGGTTCCGGAGCTGCCTCTGCTGTCCAGAAACACCTCAACTGTGAAAGAGGAGCC 3310
QY 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
DB 3311 GTGAACATTTCTGCATGTGGAGTTCTTGTCTCCAGAAAGGAAATTTCTCGCGTTCTGGGTTA 3370
QY 1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
DB 3371 ATGCCCTGTACCCATGCTCTGTGACTATTTACCAACCTTAATGCAGGAGGAGGCTTCTGC 3430
QY 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
DB 3431 CTGGCTGTCTCTTTTATGGAATACCCCATTCGCTGCTTCCAGATCCATCAGAAATGT 3490
QY 1161 SerSerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeu 1180
DB 3491 TCAAGTTTTAGTTCAACTTCTCAGCGGAGGAGGAAAGTGTGGTGGCCCCCTGCTCTCTT 3550
QY 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
DB 3551 GGACATATTAAAGAGGAGCATGAAATCAGCAGTCAAGTTTTCCGGAATGCTTCTTTAAC 3610
QY 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
DB 3611 CCTGCCACAATAGTGAACCTGCCAGCAACTTGGGGGTGTTATGTGTTGTCTCTGTCCA 3670

QY 1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
DB 3671 CTTGGATATACAGGCTTAAAGTGTGAACAGACATCGATGAGTGCACCCACCTGCCTTGC 3730
QY 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
DB 3731 CTCAACATAGTGGAGTTGTAAAGACCTAGTTGGGGAATTTCAATTTGTGAGTGCCCATCAGGT 3790
QY 1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
DB 3791 TACACAGGTGAGCGGTGTGAAGAAATATATAATGAGTGTAGCTCCAGCTCTTGTTTAAAT 3850
QY 1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
DB 3851 AAAGGAATCTGTGTGATGGTGTGGCTGCTATCGTGTGCACATGTGTGAAGGATTTGTA 3910
QY 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAla 1320
DB 3911 GGCCTGCAATTGTGAAACAGAAAGTCAATGAATGCCAGTCAAAACCATCTCTTAAATAATGCA 3970
QY 1321 ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
DB 3971 GTCTGTGAAGACAGGTTGGGGATTTCTGTGCAAAATGCCACCTGATTTTGGGTACC 4030
QY 1341 ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
DB 4031 CGATGTGGAAAGAACGTGCAATGAGTGTCTCAGTCAGCCATGCAGCAAAATGAGGCTACCTGT 4090
QY 1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
DB 4091 AAAGACGTTGCCAATAGCTTTCAGATGCTGTGTGCAGCTGGCTTTCACAGGATCACACTGT 4150
QY 1381 GluLeuAsnIleAsnGlnCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
DB 4151 GAATTCAGCATCATCAATGAATGTCAATTCATGTAGTAATAGGCCACCTGTGTGGAT 4210
QY 1401 GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
DB 4211 GAATTAATTTCAATCAGTTGTAATGTCAAGCAGGATTTTCAGGCAGAAAGGTGTGAAACA 4270
QY 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
DB 4271 GAACAGTCTACAGGCTTTAACTTGGATTTTGAAGTTTCTGSCATCTATGGATATGTCTATG 4330
QY 1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSer 1460
DB 4331 CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACCTGTACCTTCTGGAATGAAATCTCT 4390
QY 1461 AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
DB 4391 GACGACATGAACTATGGACACCCATCTCTATGCAATTAACGCGACGCGACACATACC 4450
QY 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
DB 4451 TTGCTCTGACTGATTAATAACGGCTGGGTTCTTTATGTGAATGGCAGGAGAAAGATAACA 4510
QY 1501 AsnCysProSerValAsnAspGlyArgTyrPheHisIleAlaIleThrTyrThrSerAla 1520
DB 4511 AACTGTCCCTCGGTGATGATGGCAGATGCAATCATATTGCAATCACTTGGACAGGTGCC 4570
QY 1521 AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
DB 4571 AATGGCATCTGGAAAGTCTATATATCGATGGGAATTTATCTGACGGTGTGTGGCTCTCT 4630
QY 1541 ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAsnLysLys 1560
DB 4631 GTTGGTTTCCCATACCTGCTGGTGGTGGCTGTAGTTCTGGGGCAAGAGCAGACAAAAA 4690
QY 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
DB 4691 GGAGAGGGAATTCAGCCAGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACTCTCGG 4750
QY 1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600

```

Db 4751 GACTATGTCCTGCTCCACACAGCTGAAGTCACTGGCTACTCTCTGCCAGAGAACTC 4810
Qy SerLysGlyAenValLeuAlaThrProAspPheLeuSerGlyValLeuValGlyLysVallys 1620
Db 4811 AGTAAAGAAACGTTGTAGCATGGCTGATTTCTTGTTCAGGATTTGGGGAAAGTGAAG 4870
Qy IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 4871 ATCCATTCTAAGACATATTTGTTCTGATTGATGCCACCGCTAGGAGGGTCAGTGCTCAT 4930
Qy LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 4931 CTGAGAACTGCAATCTGAAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTTCTGTGATCCA 4990
Qy GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
Db 4991 GGCTTCAGCTGGTGGGAAACCTGTGACGACTACTGTCTGATCAAGACAGTGGACACA 5050
Qy ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db 5051 CCACCTTCCTCACTGTGAACGCAATAGCTGTGGGTGCCACCTCTCTTTGGAGAAATGGCTTC 5110
Qy HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db 5111 CATTGAGCCGATGACTTCTATGCTGCGCAGCAGTAACCTACCACTGCAACAATGGCTTAC 5170
Qy TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSer 1740
Db 5171 TATCTATTGGGTGACTCAGGATGTTCTGTACGATATATGGAGCTGGACGGCGTTTCA 5230
Qy ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5231 CCATCTCGCTTGTATGTCGATGAGTGCAGTTGCAGTTGCATGATTTGATGAGCATGCTTCT 5290
Qy CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db 5291 TGCTTGAACTGAGTGAATCTTACATATGTTCTATGTGTCACCGTACACAGGATGGG 5350
Qy LysAsnCysAlaGluProIleLysCysAlaValProGlyAsnProGluAsnGlyHisSer 1800
Db 5351 AAAAATCTGTCAGAACTTATAAATGTAAGCTCCAGGAAATCCGAAATGGCCACTCC 5410
Qy SerGlyGluIleTyrThrValGlyAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db 5411 TCAGGTGAGATTTATACAGTAGTGTCCGAAAGTCAATTTTGTGTGTCAGGAGGATACCA 5470
Qy LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisLeuIlePro 1840
Db 5471 TTGATGGAGTAAACCAAAATACATGTTTGGAGTCTGGAGAAATGGAATCACTAATACCA 5530
Qy TyrCys 1842
Db 5531 TATTGT 5536
RESULT 3
ID ABT08491 standard; cDNA; 11152 BP.
AC ABT08491;
XX 28-NOV-2002 (first entry)
DT Human novel protein NOV2b coding sequence SEQ ID NO: 7.
DE Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW antiaddictive; gene; ss.

```

```

OS Homo sapiens.
PX WO200259315-A2.
XX 01-AUG-2002.
XX 19-DEC-2001; 2001WO-US050076.
XX 19-DEC-2000; 2000US-0256619P.
PR 19-JAN-2001; 2001US-0262959P.
PR 28-FEB-2001; 2001US-0272408P.
PR 20-APR-2001; 2001US-0285189P.
PR 26-JUL-2001; 2001US-0308039P.
PR 09-AUG-2001; 2001US-0311266P.
XX (CURA-) CURAGEN CORP.
PA Shimkets RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog P, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
PI Rotherne V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX WPI: 2002-666903/71.
DR P-PSDB; ABJ10590.
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.
XX Claim 8; Page 43-45; 363pp; English.
XX The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
CC graft-versus-host disease. The present sequence is a coding sequence of
CC the invention
XX SQ Sequence 11152 BP; 2951 A; 2606 C; 2711 G; 2884 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 11152
Score: 9862.00 Matches: 1781
Percent Similarity: 97.72% Conservative: 16
Best Local Similarity: 96.85% Mismatches: 40
Query Match: 96.51% Indels: 2
DB: 6 Gaps: 1
US-09-977-053-6 (1-1842) x ABT08491 (1-11152)
Qy 4 ArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThrPheGlnGln 23
Db 83 AGAATTTGCGCGGCTTGCTGGGCTCTGGCGCTCTGCTTTCGGGCTGGCGGACCTTTTCAGCAG 142
Qy 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db 143 ATGTCCCGCTCGCAATTTTCAGCTTCCGCTCTTCCCGGAGACCGCGCGGGGCCCC 202
Qy 44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
Db 203 GGGAGTATCCCGCGCGCGCGCTCTCTGCGGAGAAAGCGCGGAGAGAGTGAGGCGG 262
Qy 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83

```


Db 263 CTGGCCAGGCGTTC-----CGCGTCGGCTGCTCGGGAGCTCAGCGAGCGCCCTGGAG 316
Qy 84 LeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSerGluLeu 103
Db 317 CTTGTCTTCTCTGGTGGATGATTGCTCAGCGTGGCGGAAGTCAACTTCCGCGAGCGAGCTC 376
Qy 104 MetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db 377 ATGTTCTGTCGGCAAGCTGCTGTCGAGTTCCTCCCGTGGTCCCAAGCGCAAGCGGTGGCC 436
Qy 124 IleValThrPheSerSerLysAsnTyrValValProArgValAspTyrLysSerThrArg 143
Db 437 ATCTGTGACCTTCTCTGTCCAAGAACTACGTGGTGGCGCGGTGCAATTACATCTCCACCCGC 496
Qy 144 ArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGlnIleProAlaIleSerTyrArg 163
Db 497 CGCGCGCGCCAGCACAAAGTGGCGCTGCTCTCCAAAGAGATCCCTGCCATCTCTTACCGA 556
Qy 164 GlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
Db 557 GTGGCGGACCTTACNCCAAAGGGCGCTTCCAGCAAGCGCGCAAAATCTCTTTCATGCT 616
Qy 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAsp 203
Db 617 AGAGAAAACTCAACAAAGTGTATTCTCTCATCTGATGGATATTCCAATGGGGGAGAC 676
Qy 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db 677 CCTAGACCAATTCAGCGCTCAGTGGAGATTCAGGAGTGGAGATCTTCACTTTTGGCAT 736
Qy 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGlnHisCys 243
Db 737 TGGCAAGGAGACATTCGAGAGCTGATGATCATGGCTTCCACCCCAAGGAGGAGCACTGT 796
Qy 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeuHisGluAsp 263
Db 797 TACCTGCTACACAGTATTTGAAGAAATTTGAGGCTTTAGCTCGCGGGCAATTCGATGAAGAT 856
Qy 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeuCysAspGlu 283
Db 857 CTACCTTCTGGAGATTTTATTCAAGATGATATGGTCCACTGCTCATATCTTTGTGAGAG 916
Qy 284 GlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303
Db 917 GGCAGAGCACTGCTGTGACCGAATGGAGAGCTGCAAATGTGGGACACACACAGGCCATTTT 976
Qy 304 GluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThrAlaCys 323
Db 977 GAGTGCATCTGTGAAAGGGGTATTACGGGAAAGTCTGCAGTATGAATGCACAGCTTGC 1036
Qy 324 ProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCysIleProCys 343
Db 1037 CCATCGGGGACATACAAACCTTGAAGCTTCCACGAGGAGGAATCAGCAGTTGCATTTCCATGT 1096
Qy 344 ProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCysValCysArg 363
Db 1097 CCGATGAAATACACACCTCTCCACCTCGAGAGACATCCCTGGAAGACTGTGTCTGCAGA 1156
Qy 364 GluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
Db 1157 GAGGGATACAGGCGATCTGGCCAGACTGTGAACTTGTCTCACTGCCCTGCGCTTGAAGCCT 1216
Qy 384 ProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGly 403
Db 1217 CCGCAAAATTTGTTACTTTTATCCAAACACTTGCACCAACCACTTCAATGAGCGCTGTGG 1276
Qy 404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsn 423
Db 1277 GTCCGATGTACCTCGATTTGATCTTTGGGAGGAGCATCATCTTATGTCTACCCCAT 1336
Qy 424 GlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGln 443
Db 1337 GGTTTGTGTGTCGGTTTTCAGAGAGCTACTGCAGAGTAAAGAACATGTCTCATCTCCGCCAG 1396

Qy 444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeu 463
Db 1397 CCGAAACATGGCCACATCAGCTGCTTCTACAGGGAATGTTATATAGACACACATGTTT 1456
Qy 464 ValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGlyAsn 483
Db 1457 GTTGCTGTGATGAAGGGTACAGACTAGAAGGCAGTGAATAAGCTTACTTGTCAAGGAAC 1516
Qy 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGluMet 503
Db 1517 AGCCAGTGGGATGGGCCAGAACCCCGGTGTGTGGAGCGCCACTGTTCCACTTTCAGATG 1576
Qy 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
Db 1577 CCAAGATGTCTCATATCCCCCACTGTGGCAAGCAGCCAGCCAAATTTGGGACG 1636
Qy 524 IleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
Db 1637 ATCTGCTATGTAAGTTGCCGCCAAGGTTTCATTTTATCTGGAGTCAAAAGAAATGCTGAGA 1696
Qy 544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAspValGlu 563
Db 1697 TGTACCACCTTCTGGAAATGGAATGTCCGAGTTTCAGGACCTGTGTGTAAGACGTGGAG 1756
Qy 564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlnGlnAspSer 583
Db 1757 GCTCTCAATCACTGCTCTAAGGACATAGAGCTAAGACTCTGGAAACAGCAGATTTCT 1816
Qy 584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerVal 603
Db 1817 GCCAATGTTTACCTGGCAGATTTCCAAAGCTTAAAGACAACTCTGGTGAAGGTGTCAGTC 1876
Qy 604 HisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleVal 623
Db 1877 CGCGTTTCATCCAGCTTTCACCCCACTTACTTTTCCAAATGGAGATGTGTCTATCGTA 1936
Qy 624 TyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
Db 1937 TACACGGCAACTGACCTATCCGCAACCCAGCCAGCTGCATTTTCCATATCATACAGTTATT 1996
Qy 644 AspAlaGluProProValIleAspTyrCysArgSerProProValGlnValSerGlu 663
Db 1997 GATGCAGAACCACTGTCTATAGACTGTGTGCAGATCTCCACTCCCGTCCAGGTCTCGAG 2056
Qy 664 LysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu 683
Db 2057 AAGGTACATGCCCAAGCTGGGATGAGCTTCAGTTCTCAGACCACTCAGGGGCTGAATTG 2116
Qy 684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
Db 2117 GTCATTTACGAAGTCTATACAGAGGAGACTTTTCCCTCAGGGGAGACTATAGTACAG 2176
Qy 704 TyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLys 723
Db 2177 TATACAGCCACTGACCCCTCAGGTAAATTAACAGGATATGTGATATCCATATTTGTCTGAA 2236
Qy 724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
Db 2237 GGTTCCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATATGCACTCCAGAT 2296
Qy 744 AsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
Db 2297 AATACTGGAGTCAACTGTACATTTAACTTGTGGAGGGCTACGATTTTCAGAGGGTCT 2356
Qy 764 ThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrThrGlu 783
Db 2357 ACTGACAGGTATTATTGTGCTTATGAAGATGGCGCTCTGGAAACCAACATATATACCACTGAA 2416
Qy 784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
Db 2417 TGGCCAGACTGTGCCAAAAAAGCTTTTGGCAACACCGGGTTCAAGTCTCTTGTAGATGTTCT 2476

Qy 804 TyrLysAlaAlaArgCysAspThrAspLeuMetLysLysPheSerGluAlaPheGlu 823
Db 2477 TACAAAGCAGCTGGTTGTGATGACTCAGATCTGATGAAGAAGTTTCTGAAGCATTTGAG 2536
Qy 824 ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
Db 2537 ACGACCTCGGGAATAATGGTCCCATCATTTGTGTGTGATGAGAGGACATTTGACTGCGAGA 2596
Qy 844 LeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly 863
Db 2597 CTGGAGGAGAACCTGACCAAAAATATTTGCCTAGAATATATATGACTATGAAAATGCG 2656
Qy 864 PheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAsp 883
Db 2657 TTTGCAATTGGTCCAGTGGCTGGGCTGAGCTAATAGGCTGATTAATCTTTACGATGAC 2716
Qy 884 PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle 903
Db 2717 TTCTGGGACACTGTGCAAGAACAGCCACAGCATCGGCAATGCCAGTCTTCACGGATT 2776
Qy 904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
Db 2777 AAAAGAGTGGCCCATTTATCTGACTATTAATAATTAGTTAATTTTAAATCATCACAGTACT 2836
Qy 924 ValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGlnArgLeuLeu 943
Db 2837 GTGCCATTTACCGATGAAGAAATGATACCTTGAATGGGAAAATCAGCAACAGACTCCTT 2896
Qy 944 GlnThrLeuGluThrIleThrAsnLysLysLysArgThrLeuAsnLysAspProMetTyr 963
Db 2897 CAGACATTTGGAACATATCACAATAAATGAAAGAGCTCTCAACAAAGACCCCATGAT 2956
Qy 964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLys 983
Db 2957 TCCTTTGAGCTTGCAATGAGAAATATTAATAGCCGACAGCAATTCATTAGAACAAAG 3016
Qy 984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003
Db 3017 GCTTCCCTCTTCGACAGACAGGCTCAGTGTGTGAGAGGCGGTATGTGTCAATTTGCCCT 3076
Qy 1004 LeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr 1023
Db 3077 TTGGGAACCTATTAATAATCTGAAACATTTCACTGTGAAAGCTGCGGATCGATCCCTAT 3136
Qy 1024 GlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyr 1043
Db 3137 CAGATGAGGAAGGCACTTGAGTGCAAGCTTTGCCCTCTGCGGATGTACACGGAATAT 3196
Qy 1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
Db 3197 ATCCATTCAAGAAACATCTCTGATTTGTAAGCTCACTGTAAACAGGACCTTACTCATAC 3256
Qy 1064 SerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySer 1083
Db 3257 AGTGACTGTGAGACTGTGAAATCGTGTCACTGGGCATTTATCAACCATAAATTTGGTTCC 3316
Qy 1084 ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle 1103
Db 3317 CGGAGTCCCTCTCGTGTCCAGAAACACCTCAACTGTGTGAAAGAGGAGCGCGTGAACATT 3376
Qy 1104 SerAlaCysGlyValProCysProGluLysPheSerArgSerGlyLeuMetProCys 1123
Db 3377 TCTGCATGTGGAGTTCTCTGTCTCAGAGGAAATTTCTCGCGTCTCTGGGTAAATGCCCTGT 3436
Qy 1124 HisProCysProCysAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys 1143
Db 3437 CACCATGTCTCTGGACTATTACCACTAATGACGAGGAGGCGCTTCTGCTCGCCCTGT 3496
Qy 1144 ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerPhe 1163
Db 3497 CCTTTTATGNACTACCCCATTCGCTGGTCCAGATCCATCCATCAGAAATGTTCAAGTTT 3556
Qy 1164 SerSerThrPheSerAlaAlaGluSerValProProAlaSerLeuGlyHisIle 1183

Db 3557 AGTTCAACTTTCTCAGCGCAGAGAAAGTGTGGTCCCCCTGCTCTCTTTGGACATATT 3616
Qy 1184 LysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHis 1203
Db 3617 AAAAGAGCATGAATCAGCAGTCAAGCAAGTCAATGAATGCTTCTTTAAACCTTCCAC 3676
Qy 1204 AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr 1223
Db 3677 AATAGTGAACCTGCGCAAACTTGGGGGTGTATGTGTGTCTCTGTCCACTTGGATAT 3736
Qy 1224 ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsn 1243
Db 3737 ACAGTTTAAAGTGTGAACAGACATCATGATGATGAGCCCACTGCTTGCCTCAACAT 3796
Qy 1244 GlyValCysLysAspLeuValGlyPheIleCysGluCysProSerGlyTyrThrGly 1263
Db 3797 GAGTGTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGCCTCATCAGTTACACAGT 3856
Qy 1264 GlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIle 1283
Db 3857 AAGCACTGTGAATGAACATCAATGAATGTCAGTCTAATCCATGTAGAATCAGGCCACC 3916
Qy 1284 CysValAspGlyValAlaGlyTyrArgCysValLysGlyPheValGlyLeuHis 1303
Db 3917 TGTGTGATGAATTAATTCATACAGTTGTAAATGTACGCCAGGATTTTCAGGCAAAAG 3976
Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
Db 3977 TGTGAACAGGTATGATCACTCACTGCTTATTATTAACCTTAAATATGACAGTCTGTGA 4036
Qy 1324 AspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
Db 4037 GACCAGGTGGGGATCTTGTGCAAAATGCCCACTGGATTTTGGGTACCCGATGGA 4096
Qy 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
Db 4097 AAGAACGTGATGATGATGCTCAGTCAGCCATGCAAAAATGGAGCTACCTGTAAAGAGGT 4156
Qy 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
Db 4157 GCCAATAGCTTCAGGTGCTGTGTGAGTGCCTTACAGGATCACTGTGAATGAAC 4216
Qy 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
Db 4217 ATCAATGAATGTCACTTAATCCTATGAAATCAGGCCACCTGTGTGATGAATTAAT 4276
Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlnSer 1423
Db 4277 TCATACAGTTGTAAATGTACCCAGGATTTTCAGGCAAAAGGTGTGAACAGAACAGTCT 4336
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
Db 4337 ACAGGCTTTAACTGCAATTTTGAAGTTCCTGCACTATGATATGCTCATGTAGATGCC 4396
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspMet 1463
Db 4397 ATGCTCCCATCTCTCCATGCTTAACCTGTACCTTCTGATGAATCCCTCTGACGACATG 4456
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
Db 4457 AACATGGAACCAATCTCTATGCACTTGTATACGCGAGGACGACATACCTTGTCTCTG 4516
Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
Db 4517 ACTGATTAACCGGTGGGTCTTTATGTGAATGGCAGGAAAAGATAACAATCTGTCCC 4576
Qy 1504 SerValAsnAspGlyArgTyrHisIleAlaIleThrThrSerAlaAsnGlyIle 1523
Db 4577 TCGGTGAATGATGGCAGATGGCATATATGTCAATCCTCTGGACAGTACTGTGTGAGGCC 4636
Qy 1524 TrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu 1543

Db 4637 TGGAGGCTCATATAAATGGGGAATTTATCTGACGGTGTACTGGCTCTCCATTGGCAAA 4696
Qy 1544 ProLeuProGlyGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysGlyGluGly 1563
Db 4697 GCATACCTGGTGGGGTGCATTAGTCTTTGGGCAAGAGCAGACAAAAGAGAGGGG 4756
Qy 1564 PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal 1583
Db 4757 TTACACCGGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGACTATGTC 4816
Qy 1584 LeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGly 1603
Db 4817 CTGTCTCCACAGCAGGTGAAGTCACTGGCTACTCTCTGCCAGAGAACTCAGTAAAGGA 4876
Qy 1604 AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLysIleAspSer 1623
Db 4877 AACGTGTTAGCATGGCTCATTTCTTGTCAAGAAATGTGGGGAAGTGAAGATCGATTCT 4936
Qy 1624 LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr 1643
Db 4937 AAGAGCATATTTTGTCTGATTCCTCCACCGCTGGAGGGTCACTGCTCATCTGAGACT 4996
Qy 1644 AlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln 1663
Db 4997 GCATCTGAAGATTAAACCCAGGTTCCAAAGTCAATCTGTCTGTGAACCCAGGCTTCCAG 5056
Qy 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuPro 1683
Db 5057 CTGGTCCGGAACCTGTGCAGTACTGTCTGTAATCAAGGACAGTGGACACCAACCACTCCCC 5116
Qy 1684 HisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAla 1703
Db 5117 CACTGTGAACCATTCGCTGTGGGGTCCACCTCTCTTGGAGAAATGGCTTCCATTGAGCC 5176
Qy 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu 1723
Db 5177 GATGACTTCTATGCTGGCAGCACAGTAACTCAACCTACAGTGCACCAATGGCTACTATCTATTG 5236
Qy 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys 1743
Db 5237 GGTGACTCAAGGATGTTCTGTACAGATAATGGAGCTGGAAAGCGCTTTCACCATCTCTGC 5296
Qy 1744 LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn 1763
Db 5297 TTAGATGTCATGATGATGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 5356
Qy 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCys 1783
Db 5357 GTAGATGGATCTTACATATGTTTCATGTGTCCTCCACCGTACACAGAGATGGGAAACTGT 5416
Qy 1784 AlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu 1803
Db 5417 GCAGAACCTATATAATGTAAGGCTCCAGGAAATCCCGGAAATGCCACTCTCCTCAGGTGAG 5476
Qy 1804 IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly 1823
Db 5477 ATTATACAGTAGTCCCGAGTCAATTTCTGTACAGAGAGATACCACTGATGGGA 5536
Qy 1824 ValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuLeuProTyrCys 1842
Db 5537 GTAACCAAAATCACATGTTTGGAGTCTGGAGATGGAATCATCTAATACCATATTGT 5593
RESULT 4
ABT08490
ID ABT08490 standard; cDNA; 11158 BP.
XX
AC ABT08490;
XX
DT 28-NOV-2002 (first entry)
XX
DE Human novel protein NOV2a coding sequence SEQ ID NO: 5.
XX
KW Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;

KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
KW antidepressant; immunosuppressive; antibacterial; antiparasitic;
KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
KW antiaddictive; gene; ss.
XX Homo sapiens.
XX
XX
PH Key Location/Qualifiers
FT variation replace(717,G)
FT variation /*tag= a
FT variation replace(2303,A)
FT variation /*tag= b
FT variation replace(7348,T)
FT variation /*tag= c
FT variation replace(7370,C)
FT variation /*tag= d
FT variation replace(8665,G)
FT variation /*tag= e
FT variation replace(8827,C)
FT variation /*tag= f
FT variation replace(9018,A)
FT variation /*tag= g
FT variation replace(9551,A)
FT variation /*tag= h
FT variation replace(9790,T)
FT variation /*tag= i
FT variation replace(10025,G)
FT variation /*tag= j
XX W0200259315-A2.
XX
XX 01-AUG-2002.
XX
XX 19-DEC-2001; 2001WO-US050076.
XX
XX 19-DEC-2000; 2000US-0256619P.
PR 19-JAN-2001; 2001US-0262959P.
PR 28-FEB-2001; 2001US-0272408P.
PR 20-APR-2001; 2001US-0285189P.
PR 26-JUL-2001; 2001US-0308039P.
PR 09-AUG-2001; 2001US-0311266P.
XX
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Patturajan M, Vernet CM, Casman SJ, Malyankar U;
PI Shenoy S, Spytek KA, Gangolli E, Miller C, Boldog F, Li L;
PI Taupier RJ, Kekuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
PI Rothenberg M;
XX
XX WPI; 2002-666903/71.
DR P-PSDB; ABJ10589.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders e.g.
PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
PT or Alzheimer's disease.
XX
XX Claim 8; Page 38-41; 363pp; English.
XX
XX The present invention provides the protein and coding sequences of
CC several novel human proteins, designated NOVX. These can be used in the
CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
CC Lindau syndrome, Alzheimer's disease, stroke, tuberos sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,
CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or

QY 662 SerGluYsValHisAlaIaSerTrpAspGluProGlnPheSerAspAsnSerGlyAla 681
DB 2057 TCGGAGAAAGTATACATGCGCAAGCTGGATGAGCCTCAGTTCTCAGACAACCTCAGGTGCT 2116
QY 682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIle 701
DB 2117 GAATTGGTCATTACCAAGAGTATACACAGGAGACCTTTTCCCTCAAGGGGAGACTATA 2176
QY 702 ValGlnTrpThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
DB 2177 GTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATTGTC 2236
QY 722 IleIysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
DB 2237 ATAAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGAAATTTATATGCAC 2296
QY 742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
DB 2297 CAGATATATCTGGAGTCAACTGTACATTAACTTGTCTGGAGGCTATGATTTTCAGAA 2356
QY 762 GlySerThrAspIysTyrCysAlaTyrGluAspGlyValTrpIysProThrTyrThr 781
DB 2357 GGGTCTACTGCAAGTATTATGTGCTTATGAAGATGGCGTCTGGAACCAACATATACC 2416
QY 782 ThrGluTrpProAspCysAlaIysArgPheAlaAsnHisGlyPheIysSerPheGlu 801
DB 2417 ACTGAATGGCCAGCTGTGCGTAAGCGTTTTCGCAACCAACCGGGTTCAAGTCTCTTGAG 2476
QY 802 MetPheTyrIysAlaAlaArgCysAspAspThrAspLeuMetIysIysPheSerGluAla 821
DB 2477 ATGTTCTACAAAGCGCTCGTTGTGTATGACACAGATCTGATGAAGAATTTTCTGAAGCA 2536
QY 822 PheGluThrThrLeuGlyIysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
DB 2537 TTTGAGACGACCTCGGAAAAATGGTCCCATCATTTTGTAGTGATGACAGGACATTTGAC 2596
QY 842 CysArgLeuGluAsnLeuThrIysIysCysLeuGluTyrAsnTyrAspTyrGlu 861
DB 2597 TGCAGACTGGAGAGAACCTGACCAAAAAATATGCTCTAGAAATAATATGACTATGAA 2656
QY 862 AsnGlyPheAlaIleGlyProGlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyr 881
DB 2657 AATGGCTTTGCAATTGCTCAGGTGGCTGGGTGCGAGCTAATAGCTGGATTAATCTTAC 2716
QY 882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaIysSerSer 901
DB 2717 GATGACTTCTCGACACTGTGCAGAAACACAGCCACAGCATCGGCAATGCCAAGTCCCTCA 2776
QY 902 ArgIleIysArgSerAlaProLeuSerAspTyrIysIleIysLeuIlePheAsnIleThr 921
DB 2777 CGGATTAAGAAGAGTGGCCCATTTATCTGACTATAAAATTAAGTTAATTTTAAACATCA 2836
QY 922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnArg 941
DB 2837 GCTAGTGTGCCATTACCGATGAAGAAATGATACCTTGAATGGGAAATATCAGCAACGA 2896
QY 942 LeuLeuGlnThrLeuGluThrIleThrAsnIysLeuIysArgThrLeuAsnIysAspPro 961
DB 2897 CTCCTTCAGACATTTGAAAATATATACAAATAAACTGAAAAGGACTCTCAACAAAGACCCC 2956
QY 962 MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
DB 2957 ATGTATTCTCTTCAGCTTGATCAGAAATATCTATTATAGCCGACAGCAATTCATTAGAAACA 3016
QY 982 LysIysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
DB 3017 AAAAAGGCTTCCCTTCTGCAGACCGAGCTCAGTGTGTGAGAGGGCGTATGTGTGCAAT 3076
QY 1002 CysProLeuGlyThrTyrThrAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
DB 3077 TGGCCCTTTGGGAACCTATTATATCTGGAACATTTTACCTGTGAAAGCTGCCGGATCGGA 3136

QY 1022 SerTyrGlnAspGluGlyGlnLeuGluCysIysLeuCysProSerGlyMetTyrThr 1041
DB 3137 TCCATTCAGATGAAGAGGGCAACTTGAATGTCAGAGCTTTGCCCTCTGGGATGTACAGC 3196
QY 1042 GluTyrIleHisArgAsnIleSerAspCysIysAlaGlnCysIysGlnGlyThrTyr 1061
DB 3197 GAATATATCCATTCAAGAAACATCTCTGATTGTAAGCTCAGTGTATAACAAGGACCTTAC 3256
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProIysPhe 1081
DB 3257 TCATACAGTGTGACTGTGAGACTTGTGAATCGTGTCCACTGGGCACCTATTACGCCAAAATTT 3316
QY 1082 GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValIysArgGlyAlaVal 1101
DB 3317 GGTTCCTCCGAGCTCCCTCTCTGTGTCAGAAAAACACTCACTGTGAAAAGAGGAGCCGTG 3376
QY 1102 AsnIleSerAlaCysGlyValProCysProGluGlyIysPheSerArgSerGlyLeuMet 1121
DB 3377 AACATTTCTGCATGTGGAGTTCTTGTCCAGAGGAAAAATTTCTGCGTCTCTGGGTTAATG 3436
QY 1122 ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyIysAlaPheCysLeu 1141
DB 3437 CCCTGTCAACCATGTCTCGTGACTATTACCAACCTAATGACAGGAGGCCCTTCTGCGCTG 3496
QY 1142 AlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
DB 3497 GCCGTCCCTTTATGAACTTACCCCATTCCTGTGTTCCAGATCCATCAGACAGATGTTCA 3556
QY 1162 SerPheSerThrPheSerAlaAlaGluSerValValProProAlaSerLeuGly 1181
DB 3557 AGTTTATAGTTCAACTTTCTCAGCGGACAGGAAAGTGTGGTGCCTCCCTCTCTTGGA 3616
QY 1182 HisIleIysIysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro 1201
DB 3617 CATATTAAAGAGGACATGAATCAGCAGTCAGCAAGTCATGAATGCTTCTTTTAACCTC 3676
QY 1202 CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
DB 3677 TGCACACATAGTGNACCTGCCAGCAACTTGGGGTGTGTTANGTTTGTCTCTGTCCACTT 3736
QY 1222 GlyTyrThrGlyLeuIysCysGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
DB 3737 GGAATATACAGGTTTAAAGTGTGAAACAGACATCGATGATGTCAGCCACCTGCTTGCCTC 3796
QY 1242 AsnAsnGlyValCysIysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
DB 3797 AACAAATGGAGTTTGTAAAGACCTAGTTGGGGAAATTCATTTGTGAGTGCCTCATAGTTAC 3856
QY 1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnIys 1281
DB 3857 ACAGGTAGCACTGTGAAATTCAGATCAATGAATGTCACTAATCCATGTAGAATCAG 3916
QY 1282 GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValIysGlyPheValGly 1301
DB 3917 GCCACCTGTGTGGATGAATTAATTCATACAGTTGTAAATGTGAGTGCAGCCAGGATTTTCAG 3976
QY 1302 LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
DB 3977 AAAAGGTGTCAACAGGTATGTATCAACTCAGTGTGTTATTAATACCTTTAATATGAGCTC 4036
QY 1322 CysGluAspGlnValGlyGlyPheLeuCysIysCysProProGlyPheLeuGlyThrArg 1341
DB 4037 TGTGAGACCAAGTGTGGGGATCTTGTGCAAAATGCCCACTGGATTTTGGGTACCCGA 4096
QY 1342 CysGlyIysAsnValAspGluCysLeuSerGlnProCysIysAsnGlyAlaThrCysIys 1361
DB 4097 TGTGGAAGAACCTGCAATGTCTCAGTCAGCATGCAAAAATGGAGCTACCTGTATAA 4156
QY 1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
DB 4157 GAGGGTGCCAAATAGCTTCAGGTGCTGTGTGACGTGCTTCAGAGATCACTGTGAA 4216
QY 1382 LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401

Db 4217 TTTGAACATCAATGAATGTCAGTCTAAATCCATGTAGAAATACAGGCACCTGTGTGATGAA 4276
Qy 1402 LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
Db 4277 TTAATATTCATACAGTGTGAATGTGAGCCAGGATTTTCAGGCCAAAAGGTGTGAACAGAA 4336
Qy 1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
Db 4337 CAGTCTACAGGCTTTAACTCTGGATTTTGAAGTCTTCTGGCATCTATGGATATGTATGCTA 4396
Qy 1442 AspGlyMetLeuProSerLeuHieAlaLeuThrCysThrPheTyrMetLysSerSerAsp 1461
Db 4397 GATGGCATGCTCCCATCTCTCATGCTCTAACCTGTACCTTCTGGATGAATCTCTGAC 4456
Qy 1462 AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeu 1481
Db 4457 GACATGAATATGGAACACCAATCTCTATGAGTTGATACGGCAGCGACCAATACCTTG 4516
Qy 1482 LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsn 1501
Db 4517 CTCCTGACTGATTAATACCGGTGGTCTTTATGTGAATGGCAGGAAAAGATACCAAC 4576
Qy 1502 CysProSerValAsnAspGlyArgTyrHieIleAlaIleThrTyrThrSerAlaAsn 1521
Db 4577 TGTCCCTCGGTGAATGATGGCAGATGGCATATATTGCAATCATTGACCAAGTACTGGT 4636
Qy 1522 GlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerVal 1541
Db 4637 GGAGCCTCGAGGCTCTATTAATGGGGAATTAATCTGAGCGTGTACTGGCTCTCCATT 4696
Qy 1542 GlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLysGly 1561
Db 4697 GGCAAGGCATACCTGTGGGTGGGTGTCATTAGTCTTGGGGCAAGAGCAACACCAAAAGCA 4756
Qy 1562 GluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAsp 1581
Db 4757 GAGGGTTCACCCCGCTGAGTCTTTGTGGCTCCATPAGCCAGCTCAACTCTGGGAC 4816
Qy 1582 TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer 1601
Db 4817 TATGTCTCTGCTCCACAGCAGTGAAGTCACTGGCTACTCTCTGCTCCAGAGCAACTCAGT 4876
Qy 1602 LysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLysIle 1621
Db 4877 AAAGCAACAGTGTAGCATGGCTGATTTCTTGTTCAGGAATTTGTGGGAAAAGTGAAGATC 4936
Qy 1622 AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeu 1641
Db 4937 GATTCTAAGAGCATATTGTGTTCTGATTTGCCACCGCTTGGGAGGTCAAGTGCCTCATCTG 4996
Qy 1642 ArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly 1661
Db 4997 AGACTGCATCTGAGATTTAAACACAGGTTCCAAAGTCAATCTGTTGTGAACACGCG 5056
Qy 1662 PheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnPro 1681
Db 5057 TTCAGCTGGTGGGAAACCTCTGTGAGTACTCTCTGAATCAAGGACAGTGGACACACCA 5116
Qy 1682 LeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHis 1701
Db 5117 CTCCTCCACTGTGAACGATTTGCGTGTGGGGTGCCACCTCTCTTGGAGATGGCTTCAT 5176
Qy 1702 SerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyr 1721
Db 5177 TCAGCCGATGATCTTCTATGCTGGCAGCAGTAACTTACCAGTGCACCAATGGCTACTAT 5236
Qy 1722 LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerPro 1741
Db 5237 CTATTGGGTGATCTCAAGGATGTCTGTACAGATAATGGAGCTGGAACGGCGTTTCACCA 5296
Qy 1742 SerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHieAlaSerCys 1761

Db 5297 TCCTGCTTAGATGTGATGAGTGTGAGTGTGATTCAGATTCAGATTTGTAGTACGATCTTCTTGC 5356
Qy 1762 LeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLys 1781
Db 5357 CTGAACGTAGATGATCTCATATATTTTCATGTGTCCACCGTACACAGGAGATGGGAAA 5416
Qy 1782 AsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSer 1801
Db 5417 AACTGTGCGAAGCTATTAATGTAAGCTCAGGAAATCCGAAATTCGGCAATCTCTCA 5476
Qy 1802 GlyGluIleTyrThrValGlyValAlaValThrPheSerCysGlnGluGlyTyrGlnLeu 1821
Db 5477 GGTGAGATTTATACAGTAGGTGCGGAGTCAATTTTCTGTGTCAGGAAGTACCACTTG 5536
Qy 1822 MetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisLeuIleProTyr 1841
Db 5537 ATGGAGTAAACCAAAATCACATGTTTGGAGTCTGGAGATTCGAATCATCTAATACCATAT 5596
Qy 1842 Cys 1842
Db 5597 TGT 5599
RESULT 5
AAD32026 standard; cDNA; 11230 BP.
XX AAD32026;
AC AAD32026;
DT 18-JUN-2002 (first entry)
XX Mouse C3b/C4b complement receptor-like cDNA.
DE Mouse; C3b/C4b complement receptor-like protein; CR-like; diabetes;
XX immune system disorder; rheumatoid arthritis; psoriatic arthritis;
KW inflammatory arthritis; osteoarthritis; inflammatory joint disease;
KW autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KW transplant rejection; graft versus host disease; atherosclerosis; lupus;
KW stroke; Alzheimer's disease; ischaemic condition; neotropic; restenosis;
KW myocardial infarction; ischaemia; metabolic disorder; obesity;
KW reproductive disorder; infertility; nervous system disorder;
KW gene therapy; immunomodulatory; antipsoriatic; antiinflammatory;
KW neuroprotective; vasodilator; cardiant; anorectic; gene; ss.
XX Mus musculus.
OS
XX
FH Key Location/Qualifiers
FT CDS 95..10879
FT /*tag= a
FT /product= "Mouse C3b/C4b CR-like protein"
FT /transl_except= (pos.5363..5365, aa:Xaa)
FT /note= "Xaa corresponds to an unknown amino acid; CDS
FT does not include start codon"
FT /partial
XX
PN WO200210388-A2.
XX
PD 07-FEB-2002.
XX
PF 24-JUL-2001; 2001WO-US023548.
XX
PR 01-AUG-2000; 2000US-0222438P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Welcher AA, Elliot GS;
XX
XX WPI; 2002-257381/30.
XX P-PSDB; AAE20147.
XX
XX Nucleic acid encoding a novel C3b/C4b Complement Receptor (CR)-like
XX nucleic acid molecule, useful for treating, preventing and diagnosing
XX rheumatoid arthritis, psoriatic arthritis, inflammatory arthritis, and
XX multiple sclerosis.
PT

Db 1913 CAGCAGGACTCTGCTTAATGCCACCTGGCGAGTCCCAACAGCTTAAGACCACTCTGCTGAA 1972
 Qy LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
 Db 1973 AAGGTGTCACTCAGTCCAGCTCCAGCAGCTTTACCCCACTTACCTCTCTCCCAATTGGAGAC 2032
 Qy 620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
 Db 2033 GTGGCCATACCTACAGCGCAACCCAGCTCATCCGTAACCAAGCCAGCTGCACTTCTTAC 2092
 Qy 640 IleLysValIleAspAlaGluProValIleAspTyrCysArgSerProProVal 659
 Db 2093 ATTAAAGTCATTGATGTGGACCGCTGTCAATAGATTGTGTGCGATCTCCACCTCCATC 2152
 Qy 660 GlnValSerGluLysValHisAlaIleAspTyrAspGluProGlnPheSerAspAsnSer 679
 Db 2153 CAGGTGCTAGAGAGAGCAGCTCCAGCTGGGATGAGCTCAGTTCTCAGACCACTCC 2212
 Qy 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
 Db 2213 GGGGCTGAATTGCTCATACCAAGCAGTCCACACAGGGCGCATGTTCTCATGGGGAA 2272
 Qy 700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
 Db 2273 ACGGTGGTGTGTACACAGCCACTGACCCCTCAGGCAACACAGGACCTGTGACATCCAC 2332
 Qy 720 IleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739
 Db 2333 ATTGTTCATAAAAGGTTCTCCCTGTGAGTCCCTTCACCCCTGTAAACGGGACCTTATC 2392
 Qy 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
 Db 2393 TGTGCCAGGATAGTGTCTGGAGTAACTGTAGCTGAGCTGCAAGGAGGCTATGATTTTC 2452
 Qy 760 ThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThr 779
 Db 2453 ACAGAAAGGTCACTCAGAGAGTACTACTGTCTTTTGAAGTGTATCTGGAGACCAACA 2512
 Qy 780 TyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSer 799
 Db 2513 TACTCTACAGAAATGCCAGACTGTGTATATAAACGTTTTCGCAACCACTGTTTCAAGTCC 2572
 Qy 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
 Db 2573 TTTGAAATGCTATACAAAACCACTCGCTGTGATGACATGATCTGTTTGAAGATTTTCT 2632
 Qy 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
 Db 2633 GCAGCATTTGAGACTACCTTGGGGAACATGCTCCCTCTTTTGTAAACGATGCTGATGAC 2692
 Qy 840 IleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
 Db 2693 ATTGACTGCAGACTGGAGGAC---CTGACCAAAAATACTGCACTCAGTATATATCAAC 2749
 Qy 860 TyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyr 879
 Db 2750 TATGAAATGGCTTTTGCATTTGGACAGAGGCTGGGGTGCAGGCAACAGGCTGGATTAT 2809
 Qy 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
 Db 2810 TCTACATCATCTTCTGATGTTGTACAGAAACACCCACCGATGTGGGCAAGGCCAGA 2869
 Qy 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn 919
 Db 2870 TCGTCACGGATTAAGAAAGTGTCCCATTTGTCTGACCCCAAAATTCAGCTAATTTTAAAC 2929
 Qy 920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln 939
 Db 2930 ATCAGAGCTAGCGTGCCTCCAGAGGAAAGAAACGATACCCCTTGAATTGGAGAAATCAG 2989
 Qy 940 GlnArgLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959

Db 2990 CAGCGACTCATTTACAGCATTGGAAAACAATCAACAATCGCTGAAAAAGCACCTTTGAATAAA 3049
 Qy 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeu 979
 Db 3050 GAGCCCAATGATATCTTTTCAGCTCGCTCGAAAAACAGTGTGGTGTGACAGCAATTTCCCTC 3109
 Qy 980 GluThrLysLeuAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
 Db 3110 GAAAACAGAAAGGCTTTTCTCTCTGCAGACAGGCTCTGTGCTGAGGGGGCGCATGTGT 3169
 Qy 1000 ValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArg 1019
 Db 3170 GTCAACTGCCCTCGGAACCTTACTCTCTGAGCAATTCACCTGTGTGAAGCTGCCCTC 3229
 Qy 1020 IleGlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
 Db 3230 ATGGGATCTTACCAAGATGAAGAGGCGAGCTGGAATGCAAGCTCTGTCTCCCAAGACT 3289
 Qy 1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
 Db 3290 CAGCGGATACTCCATTCAAGAGCGTCTCTGAAATGCAAGCTCAGTGTAAAGCAAGGC 3349
 Qy 1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
 Db 3350 ACTACTCTTCCAGTGGGCTGGAGACTGCGAATCGTGTCCGCTGGGTACTTATCAACCG 3409
 Qy 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
 Db 3410 GAAATGGATCCGAGAGCTGCTCTATGCCCCAGAACCCACCAACGGTGAAGAGGA 3469
 Qy 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
 Db 3470 GCGGTGACATCTCTGCTGTGTGGAGTCCCTGCCCCAGTAGAGGAAATTTCCCGTTCTGG 3529
 Qy 1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
 Db 3530 CTAAACCCCTGCTACCTGCTGCTGCTGAGACTATTACCAACCCCAATGACGAGGAGCTCTTC 3589
 Qy 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleGlu 1159
 Db 3590 TGCCTCGCTGTGCTCTTTTATGGAACCTACAAACCATCACTCGGCCACGTCATCAGACAGAC 3649
 Qy 1160 CysSerSerPheSerThrPheSerAlaAlaGluSerValValProProAlaSer 1179
 Db 3650 TGTCTCAAGTTTACTCTTCTCTCAGCAGCAGAGAGAAAGCATAGTGCCTCGTGCCTC 3709
 Qy 1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe 1199
 Db 3710 CCTGGACATTTCCCAAGCAACAGTACGAAGTCAGCAGTCAGGTCTTTTCAAGATGCTTCTTA 3769
 Qy 1200 AsnProCysHisAsnSerGlyThrCysGlnGluLeuGlyArgGlyTyrValCysLeuCys 1219
 Db 3770 AACCCCTGCCCAACAGTGGAACTGCAACAGCTTGGGGTGGTGTATGTCTGTCTGTGC 3829
 Qy 1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
 Db 3830 CCACCTGGATACAGGCTTAAAGTGTGAACAGATATTGATGAATGATGATGCTCTGCT 3889
 Qy 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheLysCysGluCysProSer 1259
 Db 3890 TGCCTCAATGTTGGAATTTGTAGAGACCAAGTTGGGGGATTCACGTCGGAATGTTCAATG 3949
 Qy 1260 GlyTyrThrGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu 1279
 Db 3950 GGCTATTTCAGGTCAATATGTGAAGAAATATTAATGAGTGTATCTCAGGCCCTTGTCTTA 4009
 Qy 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
 Db 4010 AATAAGGAACCTGCACCTGACCGCTTGGCAAGCTTACCGCTGTACTGTGTGAAGAGTAC 4069
 Qy 1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsn 1319
 Db 4070 ATGGTGTGCTACTGTGAAACAGAGCTCAATGAATGCGCAAGTCAAGCCCTTCTTAAACAAAC 4129

XX **53**

Claim 8; SEQ ID NO 15810; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB932446 to AAB95589 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 5124 BP; 1462 A; 1106 C; 1178 G; 1378 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	512
Score:	7537.00	Matches:	1347
Percent Similarity:	99.03%	Conservative:	1
Best Local Similarity:	99.56%	Mismatches:	5
Query Match:	73.75%	Indels:	0
DB:	4	Gaps:	0

US-09-977-053-6 (1-1842) x AAH16667 (1-5124)

195	Qy	ThrAspGlyTyrSerAsnGlyGlyAspProArgProIleAlaAlaSerIleuArgAspSer	214
2	Db	ACTGATGATATTCCAATGGGGGAGACCTAGACCAATTCAGACGCTCACTCGCGAATTCA	61
215	Qy	GlyValGluIlePheThrPheGlyIleTrpGlnGlyAsnIleArgGluIleuAsnAspMet	234
62	Db	GGAGTGGNGATCTTCACCTTTGGCATATGGCAAGGGACATTCGAGAGCTGAATGCATG	121
235	Qy	AlaSerThrProLysGluGluHisCysTyrIleuLeuHisSerPheGluGluPheGluAla	254
122	Db	GCTTCCACCCCAAGAGGAGGACACTGTTCCTGCTACACAGTTCCTTTGAAGAAATTCAGGCGT	181
255	Qy	LeuAlaArgAlaLeuHisGluAspLeuProSerGlySerPheIleGluAspAspMet	274
182	Db	TTAGCTGCCCGGGCATTCGATGAGATCTACCTCTGGGAGTTCCTTTTCAGATGATATG	241
275	Qy	ValHisCysSerTyrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCys	294
242	Db	GTCCACTGCTCATATCTTTGTGTGATGAAGCGAAGGACTGCTGTGACGCAATTCGGAAGCTGC	301
295	Qy	LysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLys	314
302	Db	AAATGTGGGCACACACACAGGCCATTTTGATGTGCATCTGTGAAGAAGGGGTATTACGGGAA	361
315	Qy	GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro	334
362	Db	GGTCTGCAGTATGAATGCACAGCTTCGCCCATCGGGGACATACAAACCTTGAAGGCTCACCA	421
335	Qy	GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProGlySer	354
422	Db	GGAGGAATCAGCAGTTGCATTCCTCATGTCTGTGATGAANAATCACACCTCTCCACTGGAGGC	481
355	Qy	ThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyValInThrCysGlu	374
482	Db	ACATCCCTGAAGACTGTGCTCGACAGAGGGGATACAGGGCATCTGGCGACACTGTGTGAA	541

Qy	375	LeuValHisCysProAlaLeuLysProGluAsnGlyTrpPheIleGlnAsnThrCys	394
Db	542	CTTGTCCACCTGCCCTGCCCTGGAGCCCTCCCGAAATGGTTACTTTATFCCAACAACACTTGC	601
Qy	395	AsnAsnHisPheAsnAlaLaCysGlyValArgCysHisProGlyPheAspLeuValGly	414
Db	602	AACAACCACTTCAATGCAGCCTGTGGGGTCOGATGTCACCCCTGGATTTTGATCTTGTGGGA	661
Qy	415	SerSerIleIleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTrpCysArg	434
Db	662	AGCAGCATCATCTTATGTCTTACCCTTGGTGTGGTTCGGTTTATAGAGACTACTTCGAGA	721
Qy	435	ValArgThrCysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArg	454
Db	722	GTPAAGACATGTCTCATCTCCCGCAGCCGGAACATGGCCACATCAGCTGTTCTACAGG	781
Qy	455	GluMetLeuTrpLysThrThrCysLeuValAlaCysAspGluGlyTrpArgLeuGluGly	474
Db	782	GAATGTATATAAGACAAATGTTGGTTGGCTGTGATCAAGSGTACACACTAGAAAGGC	841
Qy	475	SerAspLysLeuThrCysGlnGlyAsnSerGluTrpAspGlyProGluProArgCysVal	494
Db	842	AGTGATTAAGCTTACTTGTCAAGGAACAGCCAGTGGGATGGGCCAGAACCCCGGTGTGTG	901
Qy	495	GluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCys	514
Db	902	GAGCGCACCTGTTCCACTTTCAGATGCCCAAGATGTCATCATATATCCCCCAACAACCTGT	961
Qy	515	GlyLysGlnProAlaLysPheGlyThrIleCysTrpValSerCysArgGlnGlyPheIle	534
Db	962	GGCAAGCAGCAGCCCAATTTGGGACGATCTGCTATGTAAATGGCCGCAAGGGTTCATT	1021
Qy	535	LeuSerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyVal	554
Db	1022	TTATCTGGAGTCAAGGAATGCTGAGATGTACCACTTCTGGAAAATGGATGTGCGAGTT	1081
Qy	555	GlnAlaAlaValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGlu	574
Db	1082	CAGCGACTGTGTGTAAGACGTGGAGGCTCTCTCAATCAACTGTCTTAGGACATAGAG	1141
Qy	575	AlaIlyThrLeuGluGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLys	594
Db	1142	GCTAAGACTCTGGAACAGCAAGATCTGCCAAATGTGTACCTGGCAGATTCACACAGCTAA	1201
Qy	595	AspAsnSerGlyGlyLysValSerValHisValHisProAlaPheThrProTrpLysLeu	614
Db	1202	GACAACTCTGGTGAAGAAGGTGTCACTCCAGCTTCATCCAGCTTTCACCCCACTTACCTT	1261
Qy	615	PheProIleGlyAspValAlaIleValTyThrAlaThrAspLeuSerGlyAsnGlnAla	634
Db	1262	TTCCCAATTCCGAGATGTGTATCGTATACACGGCAACTGACCTATCTCGGCACACCGGCC	1321
Qy	635	SerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArg	654
Db	1322	AGCTGCATTTTCATATCAAGGTATTATGATGCAGAACCACTCTGTATAGACTGGTGCAGA	1381
Qy	655	SerProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGln	674
Db	1382	TCTCCACTCCCGTCCAGGCTCTCGAGAGAGTACATGTCGCAAGCTGGGATGAGCCCTCAG	1441
Qy	675	PheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu	694
Db	1442	TTCTTCAGACAACCTCAGGGGCTGAATTTGGTCAATTACCAGAAAGTCATACCAAGGAGACCTT	1501
Qy	695	PheProGlnGlyGluThrIleValGlnTyThrAlaThrAspProSerGlyAsnAsnArg	714
Db	1502	TTCCCTCAAGGGGAGACTATAGTACAGTATACGGCCACTGACCCCTCAGGCAATAACAGG	1561
Qy	715	ThrCysAspIleHisIleValIleLysGlySerProCysGluIleProPheThrProVal	734
Db	1562	ACATGTGATATCAATATGTTCATAAAGGTTCTCCCTGTGAATTCATTACACCTGTA	1621

QY 735 AsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeu 754
DB 1622 AATGGGGATTTTATATGCACTCCAGATAAATACAGAGTCAACTGATTAACCTTGTCTG 1681
QY 755 GluGlyTyrAspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGly 774
DB 1682 GAGGGCTAGTATTCACAGAAGGCTTACTGACAAAGTATTATTGCTTATGAAGATGGC 1741
QY 775 ValTyrIysPProThrTyrThrThrGluThrProAspCysAlaIysLysArgPheAlaAsn 794
DB 1742 GTCTGGAAACCAACATATACCACTGATGATGCCAGACTGTGCCAAAACGTTTGGCAAC 1801
QY 795 HisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeu 814
DB 1802 CACGGGTTCAGTCCCTTGTAGATGTTCTACAAAGCAGCTGTTGTGTATGACACAGATCTG 1861
QY 815 MetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCys 834
DB 1862 ATCAAGAAGTTTCTGAAGCATTTGAGACGACCTGGGAAAATGCTCCCATCATTTTGT 1921
QY 835 SerAspAlaGluAspIleAspCysArgLeuGluAsnLeuThrLysLysTyrCysLeu 854
DB 1922 AGTGATCGACAGACATTCGACTGGAGACTGGAGGAGAACCTGCACAAAATATTCGCTA 1981
QY 855 GluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAla 874
DB 1982 GAATATAATTTATGACTATGAAATGCTTTGCAATTTGGACAGGCTGCTGGGGTGCAGCT 2041
QY 875 AsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSer 894
DB 2042 AATAGGCTGGATTTACTTTAGACTGACTTCTTGGACACTGTGCAAGAAACAGCCACAAGC 2101
QY 895 IleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIle 914
DB 2102 ATCGCAATGCCAAGTCTCAAGGATTTAAAGAAAGTGGCCCATTTATCTGACTATAAAAT 2161
QY 915 LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeu 934
DB 2162 AGTTTAATTTTAAACATCACAGCTAGTGTGCCATTTACCCGATGAAGAAATGATACCCCT 2221
QY 935 GluTyrGluAsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLys 954
DB 2222 GAATGGGAAATTCAGCAACGACTCTTTCAGACATTTGGAACCTATCACAATAAATGAAA 2281
QY 955 ArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAla 974
DB 2282 AGGACTCTCAACAAAGACCCCATGTATTCCTTTTCAGCTTGCATCAGAAATACTTATAGCC 2341
QY 975 AspSerAsnSerLeuGluThrLysLysAlaSerProPheCysArgProGlySerValLeu 994
DB 2342 GACAGCAATTCATTAGGAACAAAAGGCTTCCCTTCTGCAAGCCAGGCTCAGTGTCTG 2401
QY 995 ArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThr 1014
DB 2402 AGAGGGCGTATGTGTCAATTCGCCCTTTGGGAACCTATTATAATCTGGAACATTTTACC 2461
QY 1015 CysGluSerCysArgIleGlySerTyrGlnAspGluGluGlnLeuGluCysLysLeu 1034
DB 2462 TGTGAAAGCTGCCCGATCCGATCCATCAAGATGAAGAGGCAACTTGAAGTCAAGCTT 2521
QY 1035 CysProSerGlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAla 1054
DB 2522 TGCCCCCTCTGGGATGTACAGGAATATATCCATTCAGAAACATCTCTGATTTGAAGCT 2581
QY 1055 GlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeu 1074
DB 2582 CAGTGTAAACAGGACCTACTCATCGAGTGGACTTGAGACTTGTGAATGCTGTCCACTG 2641
QY 1075 GlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
DB 2642 GGCACCTATCAGCCAAAATTTGGTTCCCGAGCTGCCCTCTCGTGTCCAGAAAACACCTCA 2701
QY 1095 ThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLys 1114

DB 2702 ACTGTGAAAGAGGAGCGGTGAACATTTTGCATGTGGAGTTCCTTGTCCAGAGAGAAA 2761
QY 1115 PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProhen 1134
DB 2762 TTCTCGGTTCTGGGTTAATGCGCTGCACCCATGTCTCGTACTATTACCAACTAAT 2821
QY 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySer 1154
DB 2822 GCGGGGAAGGCCCTTCTGCCCTGGCTGTCCCTTTTATGAACTACCCCATTCCTCGTTTCC 2881
QY 1155 ArgSerIleThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSerVal 1174
DB 2882 AGATCCATCACAGATGTTCAAGTTTTAGTTCAACTTCTCAGCGGACAGAGAAAGTGTG 2941
QY 1175 ValProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPhe 1194
DB 2942 GTGCCCCCTGCTCTCTTGGACATATTAAAGAGGAGCATGAATCAGCAGTCAGGTTTTC 3001
QY 1195 HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGly 1214
DB 3002 CATGAATGCTTCTTTAAACCTTGGCACATAGTGGAACTTCCAGCAACTTGGGGCTGT 3061
QY 1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGlu 1234
DB 3062 TATGTTGTCTCTGTCCACTTGGATATACAGGCTTAAAGTGTGAAACAGACATCGATGAG 3121
QY 1235 CysSerProLeuProCysLysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIle 1254
DB 3122 TGCAGCCCACTGCTGCTCAACATGAGTGTGTAAGACCTAGTTGGGGAAATTCATT 3181
QY 1255 CysGluCysProSerGlyTyrThrGlyArgCysGluGluAsnIleAsnGluCysSer 1274
DB 3182 TGTGAGTGCCCATCAGGTTACACAGGTCAGCGGTGTGAAGAAATATAATGAGTGTAGC 3241
QY 1275 SerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThr 1294
DB 3242 TCCAGTCTTGTTTAATAAAGGAATCTGTGTGATGTTGGTGTGGCTATCGTTGCACA 3301
QY 1295 CysValLysGlyPheValGlyLysCysGluThrGluValAsnGluCysGlnSerAsn 1314
DB 3302 TGTGTGAAAGGATTTGTAGGCTGCTGTAAGACAGAGTCAATGAATGCGCAGTCAAC 3361
QY 1315 ProCysLeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysPro 1334
DB 3362 CCATGCTTAATTAATGCGTCTGTGAAGACAGGTTGGGGATTTCTGTGTCAAAATGCCA 3421
QY 1335 ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLysLeuSerGlnProCys 1354
DB 3422 CCTGGATTTTGGGTACCCGATGTGGAAGAACGTCGATGAGTGTCTCAGTCAGCCATGC 3481
QY 1355 LysAsnGlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGly 1374
DB 3482 AAAAATGGAGCTACTCTGTAAGACGCTGCAATAGCTTTCAGATGCTGTGTGTGAGTGGC 3541
QY 1375 PheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsn 1394
DB 3542 TTCACAGGATCACATGTGAATGAATCAATGAATGTCAATGTAATGTCCAGTGAAGAT 3601
QY 1395 GlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSer 1414
DB 3602 CAGCCCACTGTGTGGATGAATTAATTCATACAGTGTGAATGTCCAGCCAGGATTTTCA 3661
QY 1415 GlyGlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
DB 3662 GGCAAAAGGTGTGAAAACAGAACAGTCTACAGGCTTTAAACCTGGATTTTGAAGATTTCTGCG 3721
QY 1435 IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
DB 3722 ATCTATGATATGTATGCTAGTGTGGCATGTCTCCATCTCTCCATGCTCTAACCTGTACC 3781
QY 1455 PheTyrMetLysSerSerAspAspMetAsnTyrGlyThrProLysSerTyrAlaValAsp 1474

3782 TTCTGGATGAAATCTCTGACGACATGAACATATGGAACACCAATCTCTCTATGCACTGAT 3841
 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspThrAsnGlyTrpValLeuTrpValLeu 1494
 3842 AACGGCAGGACAAATACCTTCTCTGCTGATTAATACGGCTGGGTCTTTATGTGAAT 3901
 1495 GlyArgGluLeuValLeuThrAsnCysProSerValAsnAspGlyArgTrpHisHisLeuAla 1514
 3902 GCGAGGGAAGATGATCAACAACTGCTCTCGGTGAATGATGGCAGATGCATCATATTGCA 3961
 1515 IleThrTrpThrSerAlaAsnGlyLeuTrpValLeuTrpValLeuTrpValLeuSerAsp 1534
 3962 ATCACTTGGACAAAGTGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 4021
 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
 4022 GGTGGTGTGGCTCTCTGTGGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 4060

RESULT 7

ABZ11152
 ID ABZ11152 standard; cDNA; 3448 BP.
 XX
 AC ABZ11152;
 DT 20-JAN-2003 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 34.
 XX
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 KW cell-proliferative disorder; neurodegenerative disease; bacterial;
 KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
 KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 KW arthritis; cystostatic; immunomodulator; neotropic; neuroprotective;
 KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
 KW haemostatic; vulnerary; fungicide; antibacterial; virucide; protozoacide;
 KW antiarthritic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PP 05-MAR-2002; 2002WO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2002-759812/82.
 XX
 DR P-PSDB; ABP68935.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX
 PS Claim 1; SEQ ID NO 34; 1012pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated polynucleotide (1) comprising a
 CC nucleotide sequence selected from any of 948 sequences (ABZ1119-
 CC ABZ12066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The
 CC encoded polypeptides (ASP6902-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,

diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 arthritis, etc. Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic format
 directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3448 BP; 1014 A; 754 C; 769 G; 911 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 315e-314 Length: 3448
 Score: 5485.00 Matches: 982
 Percent Similarity: 99.90% Conservative: 1
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 53.67% Indels: 0
 DB: 6 Gaps: 0
 US-09-977-053-6 (1-1842) x ABZ11152 (1-3448)
 QY 318 TyrGluCysThrAlaCysProSerGlyThrTrpValLeuTrpValLeuTrpValLeu 337
 DB 1 TATGATGACACAGCTTGCCCATCGGGGACATACAACTGAGCCTCACCAGGGAATC 60
 QY 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
 DB 61 AGCAGTTGCATTCCTCCATGTCCTCCGATGAAATACACCTCTCCACCTGGAAGCACATCCCT 120
 QY 358 GluAspCysValCysArgGluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHis 377
 DB 121 GAAGACTGTGTCTGACAGAGAGGATACAGGCATCTGGCAGACCTGTGAACTTGTCCAC 180
 QY 378 CysProAlaLeuLysProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHis 397
 DB 181 TGCCTCTGCTGAGCCTCCGAAATGTTACTTTATCCAAACACTTGCACCAACAC 240
 QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIle 417
 DB 241 TTCAATGCAGCCTCTGGGGTCCGATGTCACCTCTGGATTTGATCTTGTGGGAAGCAGCATC 300
 QY 418 IleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTrpCysArgValArgThr 437
 DB 301 ATCTTATGCTACCCCAATGGTTGTGTCGGTTCAGAGAGCTACTGCAGAGTAGAGATAGANCA 360
 QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
 DB 361 TGTCTCTCATCTCGCCAGCCGGAACATGSCCAGATCAGCTGTTCTACAGGGGAATGTTA 420
 QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTrpArgLeuGluGlySerAspLys 477
 DB 421 TATAAGACAAACATGTTGGTTCCTGCTGATGAAGGGTACAGACTAGAGGCGAGTGAAG 480
 QY 478 LeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHis 497
 DB 481 CTTACTTGTCAAGGAACAGCCAGCTGGATGGCCAGAAACCCCGGTGTGTGGAGCCGAC 540
 QY 498 CysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGln 517
 DB 541 TGTTCACCTTTTCAGATGCCAAAGATGTCATCATATCCCCCACAACCTGTGGCAAGCAG 600
 QY 518 ProAlaLysPheGlyThrIleCysTrpValSerCysArgGlnGlyPheIleLeuSerGly 537
 DB 601 CCAGCCAAATTTGGGACGATCTGCTATGTAAGTTCGCCCAAGGGTTCATTTTATCTGGA 660
 QY 538 ValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAla 557
 DB 661 GTCAAGAAATGCTGAGATGATACCTTCTGGAAATGGAATGTCGGAGTTCAGGCAGCT 720
 QY 558 ValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThr 577
 DB 721 GGTGTAAAGACGTGGAGGCTCTCTCAATCAACTGTCCTAAGGACATAGAGGCTAAGACT 780
 QY 578 LeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSer 597

Db 781 CTGGAACAGCAAGATTCTGCAATGTTTACCTGGCAGATTCCAAACAGCTAAAGACAACTCT 840
Qy 598 GlyGluValSerValHisValHisProAlaPheThrProProTyrLeuPheProIle 617
Db 841 GGTGAAGAAGGTGTCAGTCACCGTTCATCCAGCTTTTCCACCCACCTTACCTTTTCCCAATT 900
Qy 618 GlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
Db 901 GGGAGATGTTGCTATCGTATACACGGCACTGACCTATCCGGCAACCCAGCGCAGCTGCATT 960
Qy 638 PheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProPro 657
Db 961 TTCCATATCAAGTTATTGATGCAAGAACCACTGTCTATAGACTGGTGCAGATCTCCACCT 1020
Qy 658 ProValGlnValSerGluLysValHisAlaSerTrpAspGluProGlnPheSerAsp 677
Db 1021 CCGCTCCAGGTTCGGAGAGGTACATGCCGCAAGCTGGGATGAGCTTCAGTTCTTCAGAC 1080
Qy 678 AsnSerGlyValAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697
Db 1081 AACTCAGGGGCTGAATTTGGTCATTACCGAAGTCATACACAGGAGACCTTTTCCCTCAA 1140
Qy 698 GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717
Db 1141 GGGGAGACTATAGTACAGTATACAGCCACTGACCCCTCAGCCCAATACAGCAATGTGAT 1200
Qy 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
Db 1201 ATCCATATTGTCTATAAAGGTTCTCCCTGTGAAATTCATTTCCACCTGTGAAATGGGGAT 1260
Qy 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
Db 1261 TTTATATGCACTCCAGATATATCTGGAGTCACTGTATCATTAATCTTCTGGAGGCTAT 1320
Qy 758 AspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLys 777
Db 1321 GATTTTCACAGAGGGTCTACTGCAAGTATTTATGTCCTTATGAAGATGGCGTCTGGAAA 1380
Qy 778 ProThrTyrThrThrGluTrpProAspCysAlaIleLysArgPheAlaAsnHisGlyPhe 797
Db 1381 CCAACATATACCACTGAATGGCCAGACTGTGCCAAAAAAGCTTTTGCAACACCGGGTTC 1440
Qy 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLys 817
Db 1441 AAGTCTTTTGAGATGTTCTACAAAGCAGCTCGTTGTGTATGATCACACATCTGTAGAGAAG 1500
Qy 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
Db 1501 TTTTCTGGAAGCATTTGAGACGACCTCGGGAAAAATGGTCCCATCATTTTGTAGTGATGCA 1560
Qy 838 GluAspIleAspCysArgLeuGluGluAsnLeuThrIleLysTyrCysLeuGluTyrAsn 857
Db 1561 GAGGACATTTGACTGCGAGACTGGAGGAGAACCTGACCAAAAAATATTGCCCTAGAAATATAAT 1620
Qy 858 TyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeu 877
Db 1621 TATGACTATGAATATGCTTTGCAATTTGACACAGTGGCTGGGGTGCAGCTAATAGGCTG 1680
Qy 878 AspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsn 897
Db 1681 GATTTACTCTTACGATGACTTCTCTGGACACTGTGCAAGNAACAGCCACAGCATCGGCAAT 1740
Qy 898 AlaLysSerSerArgIleLysArgSerAlaProLeuSerSerAspTyrIleLysLeuIle 917
Db 1741 GCCAGGTCTCTCAGGATTAAGAAGTGGCCCAATATCTGACTATATAAATAAGTTAAT 1800
Qy 918 PheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu 937
Db 1801 TTTAACATCACAGCTAGTGTGCCATTACCCGATGAAGAAGAAATGATATACCTTTGAATGGGAA 1860
Qy 938 AsnGlnGluArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeu 957
Db 1861 AATCAGCAAGACTCTCTCAGACATTTGGAACCTATACAAATAAATCTGAAAAAGCACTCTC 1920

Qy 958 AsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsn 977
Db 1921 AACAAAGACCCCATGATTTCTTTTCAGTTGCATCAGAAATACTTATATAGCCGACAGCAAT 1980
Qy 978 SerLeuGluThrIleLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArg 997
Db 1981 TCATTAGAAAACAAAAAGGCTTCCCTTCTGCAGACAGGCTCAGTGTCTGAGAGGGGCT 2040
Qy 998 MetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer 1017
Db 2041 ATGTGTGTCAATTTGCCCTTTTGGAACTTATATATCTTGGAACTATTCACCTTGTGAAGC 2100
Qy 1018 CysArgIleGlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSer 1037
Db 2101 TGCCGATCGGATCTATCAAGATGAAGAGGCACTTGAGTGCAAGCTTTGCCCTCT 2160
Qy 1038 GlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLys 1057
Db 2161 GGGATGTACACGGAATATATCCATTCAAGAAACATCTCTGATTTGTAAGCTCAGTGTAAA 2220
Qy 1058 GluGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyr 1077
Db 2221 CAAGGCACCTACTCATACAGTGGACTTGGAGACTTGTGAATCTGTCCACTGGGCACTTAT 2280
Qy 1078 GlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLys 1097
Db 2281 CAGCCAAATTTGGTTCCCGAGGCTGCTCTCGTGTCCAGAAACACCTCAACTGTGAAA 2340
Qy 1098 ArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArg 1117
Db 2341 AGAGGACCCGTGAACATTTCTGCAATGTGGAGTTCCTTGTCCAGAGGAAATTTCTCGCT 2400
Qy 1118 SerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLys 1137
Db 2401 TCTGGGTAAATGCCCTGTCCACCATGCTCTGTGACTATTACCAACCTTAATTCAGGGAG 2460
Qy 1138 AlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIle 1157
Db 2461 GCCTTCTGCTTGGCTGCTGCTCTTTATGGAACCTACCCCATTCGCTGGTTCAGATCCATC 2520
Qy 1158 ThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSerValValProPro 1177
Db 2521 ACAGATGTTCAAGTTTGTAGTTCAACTTTCTCAGCGGCAGAGGAAAGTGTGTGCCCT 2580
Qy 1178 AlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCys 1197
Db 2581 GCCTCTCTTGGACATATTAAAGAGGCGCATGAAATCAGCAGTCAAGTTTTCATGAATGC 2640
Qy 1198 PhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCys 1217
Db 2641 TTCTTTAACTTGGCCACATAGTGGAACTCTGCCAGCAACTTGGGCGTGGTTATGTTGT 2700
Qy 1218 LeuCysProLeuGlyTyrThrGlyLysCysGluThrAspIleAspGluCysSerPro 1237
Db 2701 CTCCTGCCACTTGGATATACAGGCTTAAAGTGTGAAGTGAACATCGATAGTGCAGCCCA 2760
Qy 1238 LeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCys 1257
Db 2761 CTGCTTGGCTTCAACAAATGGAGTTTGTAAAGACCTTAGTGTGGGAAATTCATTTGTGAGTGC 2820
Qy 1258 ProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerPro 1277
Db 2821 CCATCAGTTTACACAGTTCAGCGGTGTGAAGAAATATATATAGTGTAGTCTCAGTCT 2880
Qy 1278 CysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLys 1297
Db 2881 TGTTTAAATAAAGGAATCTGTGTTGATGGTGTGCTGCTATCGTTGCACATGTGTGAAA 2940
Qy 1298 GlyPheValGly 1301
Db 2941 GGATTTGTAGGT 2952

RESULT 8	ABK541124	ABK541124 standard; cDNA; 3128 BP.	Query Match: 46.29%	Indels: 2
ID	ABK541124		Gaps: 0	
XX	AC	ABK541124;		
XX	AC	ABK541124;		
XX	DT	05-JUN-2002 (first entry)		
XX	XX	cDNA encoding human secreted protein sequence #6.		
DE	XX	Human secreted protein; autoimmune disease; hyperproliferative disorder;		
XX	KW	cardiovascular disorder; cerebrovascular disorder; infection; cancer;		
XX	KW	nervous system disorder; ocular disorder; epithelial cell proliferation;		
XX	KW	wound healing; skin aging; sunburn; transplantation; chemotaxis;		
XX	KW	tissue regeneration; food additive; preservative; cytostatic; cardiant;		
XX	KW	antiviral; antiallergic; antiinflammatory; antibacterial; antifungal;		
XX	KW	gene; ss.		
XX	OS	Homo sapiens.		
XX	XX	WO200218412-A1.		
XX	PD	07-MAR-2002.		
XX	PF	17-JAN-2001; 2001WO-US0011384.		
XX	PR	28-AUG-2000; 2000US-0228086P.		
XX	PR	04-JAN-2001; 2001US-0259516P.		
XX	PA	(HUMA-) HUMAN GENOME SCI INC.		
XX	PI	Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;		
XX	PI	Olsen HS, Moore PA, Wei P, Ebner R, Duan RD, Shi Y, Choi GH;		
XX	PI	Piscella M, Ni J;		
XX	PI	WPI; 2002-269525/31.		
XX	DR	P-PSDB; AAU91086.		
XX	XX	Seventeen nucleic acid molecules encoding human secreted proteins, useful		
XX	PT	in the prevention, treatment and diagnosis of cancer, immune disorders,		
XX	PT	cardiovascular disorders and neurological diseases.		
XX	XX	Claim 4; Page 418-419; 505pp; English.		
XX	CC	The present invention relates to the isolation of novel human secreted		
XX	CC	proteins, and the polynucleotide sequences encoding them. The secreted		
XX	CC	proteins are useful to prevent, treat or ameliorate a medical condition		
XX	CC	in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or		
XX	CC	sheep. The secreted proteins are also useful in diagnosing a pathological		
XX	CC	condition or susceptibility to a pathological condition. Antibodies to		
XX	CC	the secreted proteins can also be used in alleviating symptoms associated		
XX	CC	with disorders and in diagnostic immunoassays e.g. radioimmunoassays or		
XX	CC	enzyme linked immunosorbent assays (ELISA). Disorders which can be		
XX	CC	diagnosed or treated include autoimmune diseases e.g. rheumatoid		
XX	CC	arthritis, hyperproliferative disorders e.g. cancer, cardiovascular		
XX	CC	disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral		
XX	CC	ischaemia, angiogenesis, nervous system disorders e.g. Parkinson's		
XX	CC	disease, infections caused by bacteria, viruses and fungi and ocular		
XX	CC	disorders e.g. corneal infection. The polypeptides can also be used to		
XX	CC	aid wound healing and epithelial cell proliferation, to prevent skin		
XX	CC	aging due to sunburn, to maintain organs before transplantation, for		
XX	CC	supporting cell culture of primary tissues, to regenerate tissues and		
XX	CC	chemotaxis. The polypeptides can also be used as a food additive or		
XX	CC	preservative to increase or decrease storage capabilities. ABK54119-		
XX	CC	ABK541167 encode human secreted protein sequences		
XX	SQ	Sequence 3128 BP; 832 A; 785 C; 763 G; 746 T; 0 U; 2 Other;		
XX	XX	Alignment Scores:		
XX	Pred. No.:	1,06e-269 Length: 3128		
XX	Score:	4730.00 Matches: 863		
XX	Percent Similarity:	99.65% Conservative: 2		
XX	Best Local Similarity:	99.42% Mismatches: 2		

Alignment Scores:		
Pred. No.:	1,068-269	Length:
Score:	4730.00	Matches:
Percent Similarity:	99.65%	Conservative:
Best Local Similarity:	99.43%	Mismatches:
		2

QY 340 ePleProCysProAspGluAenHisThrSerProGlySerThrSerProGluAspCys 360
 DB 1205 CATTCCATGCTCTGATGAATCAACACCTCTCCACTGAGACACATCCCTCTGAGACTG 1264
 QY 360 sValCysArgGluGlyTyArgAlaSerGlyGlnThrCysGluLeuValHisCysProAl 380
 DB 1265 TGTCTGCAGAGGGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTCCACTGCCCTGC 1324
 QY 380 aLeuLysProGluAenGlyTyPheIleGlnAenThrCysAenAenHisPheAenAl 400
 DB 1325 CCTGAAGCCTCCCGAAATGGTTACTTTATCCAAACACCTTGCAACAACCACTTCAATGC 1384
 QY 400 aAlaCysGlyValArgCysHisPheProGlyPheAspLeuValGlySerSerIleIleLeuCy 420
 DB 1385 AGCCTGTGGGGTCCGATGTCACCTCGGATTTGATCTTGTTGGAGACGACGATCATCTTATG 1444
 QY 420 sLeuProAenGlyLeuTyrSerGlySerGlySerThrTyArgValArgThrCysProHi 440
 DB 1445 TCTACCCCAATGGTTGTGGTCCGTTCCAGAGCTACTGCAGAGTAAGAACAATGTCCTCA 1504
 QY 440 sLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyriLysTh 460
 DB 1505 TCTCCGCGAGCGCAACATGCGCACATCAGCTGTTCTACAGGGAAATGTTATATAAGAC 1564
 QY 460 rThrCysLeuValAlaCysAspGluGlyTyArgLeuGluGlySerAspLysLeuThrCy 480
 DB 1565 AACATGTTGGTTGGCTGTGATGAAGGGTACAGACTAGAAAGCGAGTAAGCTTACTTGT 1624
 QY 480 eGlnGlyAenSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerTh 500
 DB 1625 TCAAGAAACACGCGATGGGTGGGTCGAGAACCCCGGTGTGTGGAGCGCCACTGTTCAC 1684
 QY 500 rPheGlnMetProLysAspValIleIleSerProHisAenCysGlyLysGlnProAlaly 520
 DB 1685 CTTTCAGATGCCCAAGATATCATCATATATATATATATATATATATATATATATATAT 1744
 QY 520 sPheGlyThrIleCysTyValSerCysArgGlnGlyPheIleLeuSerGlyValLysG 540
 DB 1745 ATTTGGGAGCATCTGCTATGTAGTTGGCGCCAGGGTCAATTTATCTGGAGTCAAGA 1804
 QY 540 uMetLeuArgCysThrThrSerGlyLysTrpAenValGlyValGlnAlaAlaValCysLy 560
 DB 1805 AATGCTGAGATGACCACTTCTGGAATGGATGTGGAGTTCAGGAGCTGTGTGTAA 1864
 QY 560 sAepValGluAlaProGlnIleAenCysProLysAspIleGluAlaLysThrLeuGlu 580
 DB 1865 AGACGTGGAGGCTCTCAATCAATCACTGCTTAAGGACATAGAGCTTAAGACTCTGGAACA 1924
 QY 580 nGlnAspSerAlaAenValThrTrpGlnIleProThrAlaLysAspAenSerGlyGly 600
 DB 1925 GCNAGATTCTGCCAATGTTTACCTGGCAGATTCACACAGCTAAAGACAATCTGTGGTGA 1984
 QY 600 sValSerValHisValHisProAlaPheThrProProTyLeuPheProIleGlyAspVa 620
 DB 1985 GGTGTGAGTCCAGCTTCATCCAGCTTTTCCACCACTTACCTTTTCCCAATGGAGATGT 2044
 QY 620 lAlaIleValTyThrAlaThrAspLeuSerGlyAenGlnAlaSerCysIlePheHisI 640
 DB 2045 TGTCTATGTTATACAGGCAACTGACCTATCCGCAACACAGGCGAGCTGCAATTTTCCATAT 2104
 QY 640 eLysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValG 660
 DB 2105 CAAGGTTATTTGATGCAAGAACCACTGTGATAGACTGGTGCAGATCTCCACCTCCCGTCCA 2164
 QY 660 nValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAenSerG 680
 DB 2165 GGTCTCGAGAGATACATGCGCCAGCTGGGATGAGCTTCTCTCAGACAATCAAG 2224
 QY 680 yAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluTh 700
 DB 2225 GGTCTGAATTTGGTCAATACAGAGTATACACAGGAGAGCTTTTCCCTCAAGGGGAGAC 2284
 QY 700 rIleValGlnTyThrAlaThrAspProSerGlyAenAenArgThrCysAspIleHisI 720

DB 2285 TATAGTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATAT 2344
 QY 720 eValIleLysGlySerProCysGluIleProPheThrProValAenGlyAspPheIleCy 740
 DB 2345 TGTCTATAAAGGTTCTCCCTGTGAATTCATTTACACACCTGTAAATGGGGATTTTATATG 2404
 QY 740 sThrProAspAenThrGlyValAenCysThrLeuThrCysLeuGluGlyTyArgPheTh 760
 DB 2405 CACTCCAGATTAATCTGGAGTCACTGTACATTAATCTTCTTGGAGGCTATGATTTTCAC 2464
 QY 760 rGluGlySerThrAspLysTyTyCysAlaTyArgGluAspGlyValTrpLysProThrTy 780
 DB 2465 AGAAGGCTCTACTGACCAAGTATTTATGTCTTATGAAGATGGCTCTCGAAACCAACATA 2524
 QY 780 rThrThrGluTrpProAspCysAlaLysLysArgPheAlaAenHisGlyPheLysSerPh 800
 DB 2525 TACCACTGAATGGCCAGACTGTGCCAAAACCGTTTTGCAACACACGGTTCAGTCCCT 2584
 QY 800 eGluMetPheTyLeuAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerG 820
 DB 2585 TGAGATGTTCTACAAAGCAGCTCGTTGTGATGATACACAGATCTGATGAAGAAGTTTCTGA 2644
 QY 820 uAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspI 840
 DB 2645 AGCATTTGAGACGACCCCTGGGAAAATGGTCCCATCATTTTGTAGTATGCAGAGGACAT 2704
 QY 840 eAepCysArgLeuGluGluAenLeuThrLysLysTyCysLeuGluTyArgPheThrAspTy 860
 DB 2705 TGACTGCAGACTGGAGGAGAACCTGACCAAAAATATTTGCTAGATATATATATATATAT 2764
 QY 860 rGluAenGlyPheAlaIleGly 867
 DB 2765 TGAATAATGGCTTTGCAATTTGGT 2786

RESULT 9
 ABK35664/c
 ID ABK35664 standard; cDNA; 2064 BP.
 XX
 AC ABK35664;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE cDNA sequence #55 encoding novel human secreted protein.
 XX
 KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
 KW immune deficiency disorder; blood disorder; inflammatory disorder;
 KW infectious disorder; allergic condition; neurodegenerative disorder;
 KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
 KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200177289-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US010232.
 XX
 PR 06-APR-2000; 2000US-0195605P.
 XX
 XX (GEMY) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;
 PI Clark HF, Pechtel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
 XX WPI; 2002-179322/23.
 XX
 XX Six hundred and twenty three polynucleotides derived from a variety of
 XX human tissue sources which encode secreted proteins, useful for treating
 XX immune deficiencies and disorders such as autoimmune disorders.

Claim 1; Page 102; 393pp; English.

The present invention relates to the isolation of novel cDNA sequences which encode human secreted proteins. The cDNA sequences have been derived from a variety of human tissues. The invention also provides a method for producing proteins from these polynucleotide sequences. The proteins are useful for identifying compounds that modulate their activity and production. The sequences of the invention are useful for treating diseases such as hyperproliferative disorders (e.g. cancer), immune deficiency disorders (e.g. severe combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis), infectious disorders (e.g. hepatitis), allergic conditions (e.g. asthma), neurodegenerative disorders (e.g. Alzheimer's disease), liver fibrosis, coagulation disorders (e.g. haemophilia), and tumours. The polynucleotide sequences of the invention are also useful in gene therapy. ABK35610-ABK36232 represent the cDNA sequences of the invention that encode for novel human secreted proteins

Sequence 2064 BP: 512 A; 444 C; 436 G; 672 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.24e-175	Length:	2064
Fragment Score:	3133.00	Matches:	557
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	30.66%	Gaps:	0
DB:	6		

US-09-977-053-6 (1-1842) x ABK35664 (1-2064)

Qy	311	TyrTyrGlyIysGlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysPro	330
Db	2042	TAATTACGGGAAGGCTCTGCAGTAGTAATGAAATGCACAGCTTGCCCATCGGGAATACATAAAACCT	1983
Qy	331	GlulGlySerProGlyIysLeuSerSerCysIleProCysProAspGluAsnHisThrSer	350
Db	1982	GAAGGCTCACACAGGAGGAGATCAGCAGTTGCATTCCATGTCGTGATGAATATCACACCTCT	1923
Qy	351	ProProGlySerThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGly	370
Db	1922	CCACCTGGAAAGCACATCCCTCTGAAGACTGTGTCTGTCAGAGAGGGGATACAGGGCATCTGGC	1863
Qy	371	GlnThrCysGluLeuValHisCysProAlaLeuIysProProGluAsnGlyTyrPheIle	390
Db	1862	CAGACCTGTGAACCTGTTCACCTGCCCTGCCCTGAGCCCTCTGAAATAGTTACTTTATC	1803
Qy	391	GlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPhe	410
Db	1802	CAAAACACTTGCACAACCACTTCAATGCAGACCTGTGGGGTCCGATGTCACCCCTGGATTT	1743
Qy	411	AspLeuValGlySerSerIleIleLeuCysAlenProAsnGlyLeuTpsSerGlySerGlu	430
Db	1742	GATCTTGTGGAGCAGACATCACTTATGCTACCCATNGGTTGTGGTTCGGTTCAGAG	1683
Qy	431	SerTyrCysArgValArgThrCysProHisLeuArgGlnProIysHisGlyHisIleSer	450
Db	1682	AGCTACTGCAGAGTAAGAACATGTCTCATCTCCGCCAGCCGAACATGGCCACATCAGC	1623
Qy	451	CysSerThrArgGluMetLeuTyrLysThrThrCysLeuValAlaCysAspGluGlyTyr	470
Db	1622	TGTTCTACAGGGGAATGTTATATAGACAAACATGTTTGGTTGCTCTGTGATGAAGGGTAC	1563
Qy	471	ArgLeuGluGlySerAspLysLeuThrCysGlnGlyAsnSerGlnTpsAspGlyProGlu	490
Db	1562	AGACTAGAGGCAGTGTATAGCTTACTTGTTCAGAGAAACGCCAGTGGGATGGGCCAGAA	1503
Qy	491	ProArgCysValGluArgHisCysSerThrPheGlnMetProIysAspValIleIleSer	510
Db	1502	CCCCGGTGTGTGGAGCGCCACTGTTTCCACCTTTCAGATGCCCAAGAGTGCATCATATCC	1443
Qy	511	ProHisAsnCysGlyLysGlnProAlaIysPheGlyThrIleCysbTyrValSerCysArg	530

RESULT 10
AAH16567
ID AAH16

AAH16567
ID AAH16567 standard: cDNA: 1969 BP.

1442	CCCCAACACTGTGGCAGCGCCAGCCAGCAAAATTTGGGAGCGATCTGCTATGTAAGTTGGCGC	138
531	GInGlyPheIleLeuSerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTrp	550
1382	CAAGGGTTCAATTTATCTGGAGTCAAAAGAAATGCTGAGATGTACCACTTCTGGAAAAATGG	1323
551	AsnValGlyValGlnAlaValCysLysAspValGluAlaProGlnIleAsnCysPro	570
1322	AAATGTGGAGTTTCAGGCAGCTGTGTGTAAAGACGTGGAGGCTCTCTCAATCAACTGTCTCT	1263
571	LYASpIleGluAlaLysThrLeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIle	590
1262	AAGCACAATAGAGGCTTAGACTCTGGAAACAGCAGAAATTCGCCAATGTTACTTGGCAGATT	1203
591	ProThrAlaLysAspAsnSerGlyGluLysValSerValHisValHisProAlaPheThr	610
1202	CCAAACGCTAAAGACCAACTCTCGTGTGAAAAGGTGTCACTCCACGTTCTATCCAGCTTTCCACC	1143
611	ProProTyrLeuPheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSer	630
1142	CCACCTTACCTTTTCCCAATTTGGAGATGTTGCTATCGTATATACGGCAACTGACCTATACC	1083
631	GlyAsnGlnAlaSerCysIlePheHisIleLysValIleAspAlaGluProProValIle	650
1082	GGCAACCGACGCCAGCTGCAATTTTCCATATCAAGSTTATTGATGTCAGAAACCACTGTCTATA	1023
651	AspTrpCysArgSerProProValGlnValSerGluLysValHisAlaAlaSerTrp	670
1022	GACTGTGTGCAGATCTCCACCTCCCGTCCAGTCTCGGAGAGGTACATGCGCAAGCTGG	963
671	AspGluProGlnPheSerAspAsnSerGlyValaGluLeuValIleThrArgSerHisThr	690
962	GATGAGCCTCAGTTCTCAGCAACTCTCAGGGCTGGAATGGTCAATTACAGAGACTCATACA	903
691	GlnGlyAspLeuPheProGlnGlyGluThrIleValGlnTyrThrAlaThrAspProSer	710
902	CAAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAGTATACAGCCACTGACCCCTCA	843
711	GlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySerProCysGluIlePro	730
842	GGCAATAACAGGACATGTGTATATCCATATTTGTATATAAAGGTTCTCCCTGTGAAATTTCCA	783
731	PheThrProValAsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCysThr	750
782	TTTACACCTGTAAATGGGAAATTTTATATGCACTCCAGATTAATACTGGAGTCAACTGTACA	723
751	LeuThrCysLeuGluGlyTyrAspPheThrGluGlySerThrAspLysTyrTyrCysAla	770
722	TTAACTTGTCTGGAGGGCTATGATTTTACAGAAGGGTCTACTGACAAAGTATATTGTGTCT	663
771	TyrGluAspGlyValTrpLysProThrTyrThrThrGluTrpProAspCysAlaLysLys	790
662	TATGAAGATGGCGTCTGGAACCAACACATATACCACTGAATGGCCAGACTGTGCCAATAA	603
791	ArgPheAlaAsnHisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAsp	810
602	CGTTTTGCAAAACCAACCGSGTTCAAGTCTCTTGGAGATGTTCTACAAAGCAGCTCGTTGTGAT	543
811	AspThrAspLeuMetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMetVal	830
542	GACACAGATCTGATGAAGAAGATTTTCTGAAGCAATTTGAGACGACCTTGGGAAAAATGGTC	483
831	ProSerPheCysSerAspAlaGluAspIleAspCysArgLeuGluAlaAsnLeuThrLys	850
482	CCATCATTTGTGTGTATGACAGAGGACATTTGATCTGCAGACTGGAGGAGAGCACTGACCCAA	423
851	LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGly	867
422	AAATATTTGCTAGATATATATATGACTATGAAAAATGGCTTTGCAATTTGGT	372

XX AC AAH16567;
 XX 26-JUN-2001 (first entry)
 XX DE Human cDNA sequence SEQ ID NO:15640.
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX OS Homo sapiens.
 XX PN BP1074617-A2.
 XX PD 07-FEB-2001.
 XX PF 28-JUL-2000; 2000BP-00116126.
 XX PR 29-JUL-1999; 99JP-00248036.
 XX PR 27-AUG-1999; 99JP-00300253.
 XX PR 11-JAN-2000; 2000JP-00118776.
 XX PR 02-MAY-2000; 2000JP-00183767.
 XX PR 09-JUN-2000; 2000JP-00241899.
 XX PA (HELI-) HELIX RES INST.
 XX PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX DR WPI; 2001-318749/34.
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX PS Claim 8; SEQ ID NO 15640; 2537pp + Sequence Listing; English.
 XX CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 XX SQ Sequence 1969 BP; 603 A; 434 C; 430 G; 502 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 4,07e-172 Length: 1969
 Score: 3073.00 Matches: 547
 Percent Similarity: 99.45% Conservative: 0
 Best Local Similarity: 99.45% Mismatches: 3
 Query Match: 30.07% Indels: 0
 DB: 4 Gaps: 0
 US-09-977-053-6 (1-1842) x AAH16567 (1-1969)
 QY 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIle 337

DB 1 TATGATGCACAGCTTGCCCATTCGGGACATACAACTGAGCCTCACCGAGGAGATC 60
 QY 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
 DB 61 AGCAGTTGCATTCCATGTCGATGAAATCACAACCTCTCCACCTGGAAGCACATCCCT 120
 QY 358 GluAspCysValCysArgGluGlyTyrArgAlaSerDlyGlnThrCysGluLeuValHis 377
 DB 121 GAAGACTGTGCTGCAGAGAGGGATACAGGGCATCTGGCCAGACTGTGAACCTTGTCAC 180
 QY 378 CysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHis 397
 DB 181 TGCCTGCTGCTGAGCTCCGGAATGGTTACTTTATCCAAACACTTGCACACACAC 240
 QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIle 417
 DB 241 TTCAATGCACCTGTGGGTCCGATGTCACCTCGATTGATCTGTGGGAAGCAGCATC 300
 QY 418 IleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThr 437
 DB 301 ATCTTATGTC7ACCCCAATGGTTGTGGTCCGGTTTCAGAGAGCTACTGCAGAGTAAGAACA 360
 QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
 DB 361 TGTCCTCACTCTCCGACGCGAACAACATGGCCACATCAGCTGTCTACAGGGAATGTTA 420
 QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
 DB 421 TATAAGACAACATGTTTGGTGTGCTGTGATGAAGGGTACAGACTAGGAAGGCAGTATAAG 480
 QY 478 LeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHis 497
 DB 481 CTTACTTGTCAAGGNAACAGCCAGTGGATGGGCGCAGAACCCCGGTGTGTGAGCGCCAC 540
 QY 498 CysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGln 517
 DB 541 TGTTCACCTTTTCAGATGCCAAGAGTGTCAATCATATCCCCCACAACCTGTGGCAGCAG 600
 QY 518 ProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGly 537
 DB 601 CCAGCCAAATTTGGGACGATCTGCTATGTAAAGTTGCCGCAAGGGTTCATTTATCTGGA 660
 QY 538 ValLysGluMetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAla 557
 DB 661 GTCAAGAAATGCTGAGATGTACCACTTCTGGAATAATGGAATGTCGAGTTTCAGCGAGCT 720
 QY 558 ValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThr 577
 DB 721 GTGTGTAAAGCTGGAGGCTCTCTCAATCACTGTCTTAAGGACATAGAGGCTTAAGACT 780
 QY 578 LeuGluGlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSer 597
 DB 781 CTGGAACAGCAGAGATTCTGCAATGTTTACCTGGCAGATTCCACAGCTAAAGACAACCTCT 840
 QY 598 GlyGluLysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIle 617
 DB 841 GGTGAAAGGTGTGTCAGTCGCGTTCATCCAGCTTTTCAACCCACCTTACCTTTTCCCAAT 900
 QY 618 GlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
 DB 901 GGAGATGTTGCTATCGTATACAGCGCACTGACTATCCGCGCAACACAGGCGCAGCTGCATT 960
 QY 638 PheHisIleLysValIleAspAlaGluProProValIleAspTyrCysArgSerProPro 657
 DB 961 TTCCATATCAGGTTATTGATGTCAGAAACCACTGTCTATAGACTGGTGCAGACTCTCCACCT 1020
 QY 658 ProValGlnValSerGluLysValHisAlaAlaSerThrAspGluProGlnPheSerAsp 677
 DB 1021 CCGGTCCAGGCTCTCGAGAGAGGTACATGCGCAGCTGGATGAGCTCTAGTCTCTAGAC 1080
 QY 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697

Db 1081 AACTCAGGGCTGATTTGGTCAATACACAGAGTCAATACACAGAGACCTTTTCCCTCAA 1140
Qy 698 GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717
Db 1141 GGGGAGACTATAGTACAGTATACAGCCACAGCCCTCAGGCAATAACAGGACATGTGAT 1200
Qy 718 IleHisIleValIleGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
Db 1201 ATCCATATTGCTATAAAGGTTCTCCCTGTGAATCCCATTCACACCTGTAAATGGGGAT 1260
Qy 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
Db 1261 TTTATATGCACTCCAGATAATACCTGGAGTCACTGTATCACTTAATCTTGGAGGGCTAT 1320
Qy 758 AspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLys 777
Db 1321 GATTTTCACAGAGGGTCTACTGACAGATATATTGTGCTTATGAAGATGGCGTCTGGAAA 1380
Qy 778 ProThrTyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPhe 797
Db 1381 CCACATATATACCACTGAATGCCAGACTGTGCCAAAACGTTTTCGAACCCACGGGTTC 1440
Qy 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLys 817
Db 1441 AAGTCTTTGAGAGTTCTACAGACAGCTGCTGTGTATGACACAGATCTGATGAGAAG 1500
Qy 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
Db 1501 TTTTCTGAGCATTTGGAGAGACCCCTGGAAAAAATGGTCCCATCATTTTGTAGTATGCA 1560
Qy 838 GluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsn 857
Db 1561 GAGGACATTCAGCTGCAGACTGGAGAGAACCTGCACAAAATAATATGCTAGATAATAAT 1620
Qy 858 TyrAspTyrGluAsnGlyPheAlaIleGly 867
Db 1621 TATGACTATGAAAATGGCTTTGCAATTGGT 1650
RESULT 11
ABK54146
XX ID ABK54146 standard; cDNA; 1696 BP.
AC ABK54146;
XX
XX
XX 05-JUN-2002 (first entry)
XX cDNA encoding human secreted protein sequence #28.
XX Human secreted protein; autoimmune disease; hyperproliferative disorder;
KW cardiovascular disorder; cerebrovascular disorder; infection; cancer;
KW nervous system disorder; ocular disorder; epithelial cell proliferation;
KW wound healing; skin aging; sunburn; transplantation; chemotaxis;
KW tissue regeneration; food additive; preservative; cytostatic; cardiant;
KW antiviral; antiallergic; antiinflammatory; antibacterial; antifungal;
KW gene; ss.
XX Homo sapiens.
XX WO200218412-A1.
XX
XX 07-MAR-2002.
XX
XX 17-JAN-2001; 2001WO-US001384.
XX
XX 28-AUG-2000; 2000US-0228086P.
XX 04-JAN-2001; 2001US-0259516P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Komatsoulis GA, Baker KP, Birse CS, Soppet DR;
PI Olsen HS, Moore PA, Wei P, Ebner R, Duan RD, Shi Y, Choi GH;
PI Fisciella M, Ni J;
XX

WPI; 2002-269525/31.
P-ESDB; AAU91108.

Seventeen nucleic acid molecules encoding human secreted proteins, useful in the prevention, treatment and diagnosis of cancer, immune disorders, cardiovascular disorders and neurological diseases.

Claim 4; Page 435-436; 505pp; English.

The present invention relates to the isolation of novel human secreted proteins, and the polynucleotide sequences encoding them. The secreted proteins are useful to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The secreted proteins are also useful in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the secreted proteins can also be used in alleviating symptoms associated with disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which can be diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cancer, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Parkinson's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. ABK54119-ABK54167 encode human secreted protein sequences

Sequence 1696 BP; 349 A; 489 C; 475 G; 369 T; 0 U; 14 Other;

Alignment Scores:

Pred. No.:	4e-137	Length:	1696
Score:	2479.00	Matches:	464
Percent Similarity:	97.68%	Conservative:	0
Best Local Similarity:	97.88%	Mismatches:	8
Query Match:	24.26%	Indels:	4
DB:	6	Gaps:	0

US-09-977-053-6 (1-1842) x ABK54146 (1-1696)

Qy	1	MetTrpProArgLeuAlaPheCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr	20
Db	238	ATGTGGCCCTCGCCCTGGCCCTTTGTGTCTGGGGCTCTGGCGCTCTGTTCCGGCTGGCGGAC	297
Qy	21	PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro	40
Db	298	TTTCAGCAGATGTCCCGTCCGCAATTTTCAGTTCCGCTCTTCCCGAGACCGCGCS	357
Qy	41	GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg	60
Db	358	GGGGCCCCCGGAGTATCCCGCGCCCGCTCTCTGGCGCAGAAAGCGCGGGAGCAGA	417
Qy	61	ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu	80
Db	418	GTGAGCGGCTGGGGCCAGCGCTTCGGGACGCGTCCGCGCTCTCGGGAGACTCARGAG	477
Qy	81	ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValValAsnPheArg	100
Db	478	CGCCTGGAGCTTGTCTCTCTGTGATGATTCAGCTCCAGCGTGGCGAGTCAACTTCGCG	537
Qy	101	SerGluLeuMetPheValArgLysLeuSerAspPheProValValProThrAlaThr	120
Db	538	AGCGAGCTCATGTTCTCGCAAGCTGCTGTCCGACTTCCCGTGGTCCCGGCGCAG	597
Qy	121	ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle	140
Db	598	CGCGTGGCCATCGTGACCTTCTCGTCCAAAGAACTACGTGGTGGTCCCGCGCTCGATTACATC	657
Qy	141	SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluLeuProAlaIle	160


```
Qy 865 AlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPhe 884
Db 63 GCGATGGGACAGGGGGCTGGGGCGGAGCTCATAGGCTGGAGTACTCTTACGATGACTTG 122
Qy 885 LeuAspThrValGlnGlnThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLys 904
Db 123 CTGAGCCTGAGCAAGAAACAGCCACCGCATCGCATCGCATCTCAAGTACTCAAGAGCTAT 182
Qy 905 ArgSerAla - ProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerVa 924
Db 183 AAGAGGAGCGCCATTAATCTGACTATATAAATTAAGTTAATTTTAAACATCACAGCTAGTGT 242
Qy 924 lProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGlnArgLeuLeuG 944
Db 243 GCCATTACCCGATGAAGAAATGATACCTCTTGAATGGGAAATCAGCAACGACTCTTCA 302
Qy 944 nThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSe 964
Db 303 GACATTTGGAAACTATCACAAATAACTGAAAGGACTCTCAACAAAGACCCCATGTATTTC 362
Qy 964 rPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAl 984
Db 363 CTTTTCAGCTTGCAATCAAGAACTATTATAGCCGACAGCAATTCATTAGAAACAAAAAGGC 422
Qy 984 aSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLe 1004
Db 423 TTCCCCCTCTTCGACAGCAGGCTCAGTCTGAGAGGGCGTATGTGTGTCAATTCGCCCTTT 482
Qy 1004 uGlyThrTyrThrAsnLeuGluHlePheThrCysGluSerCysArgIleGlySerTyrGl 1024
Db 483 GGGAACTTATTATATCTGGAAACATTTCCCTGTGAAAGCTGCCGATCGGATCTCTATCA 542
Qy 1024 nAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIl 1044
Db 543 AGATGAGAAGGCAACTTGAGTGCAGACTTTGCCCCCTCTGGCATGTACACGGAATATAT 602
Qy 1044 eHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyrSe 1064
Db 603 CCATTCAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAGGACCTTACTCATACAG 662
Qy 1064 rGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySerAr 1084
Db 663 TGGACTTGAGACTGTGATCGTGTCTCACTGGGCACTTATCAGCCAAATTTGGTTCCCG 722
Qy 1084 gSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSe 1104
Db 723 GAGCTGCTCTCTGCTCCAGAAACACCTCAACTGTGAAAGAGGAGCGCGTGAACATTTTC 782
Qy 1104 rAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHi 1124
Db 783 TGCATGTGGAGTTCTTGTCCAGAAAGAAATTTCTGGCGTTCTGGGTTAATGCCCTGTCA 842
Qy 1124 sProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysPr 1144
Db 843 CCCATGCTCTGTGACTATTAACCACTATGTCAGGGAAGGCTTCTGCTGGCTGCTGCC 902
Qy 1144 oPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSe 1164
Db 903 CTTTTATGGAATCACTCCCACTGCTGGTTCAGATCCATCCATCAGAAATGTTCAAGTTTAT 962
Qy 1164 rSerThrPheSerAlaAlaGluLysSerValValProProAlaSerLeuGlyHleIleLy 1184
Db 963 TTCAACTTTCTCAGCGGAGAGGAAAGTGTGGTGGCCCCCTGCTCTCTTGGACATATTAA 1022
Qy 1184 sLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisAs 1204
Db 1023 AAAGAGGAGTGAATACAGCATGACAGTTTCATGATGCTCTCTTAACCTTGGCCACA 1082
Qy 1204 nSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrTh 1224
Db 1083 TAGTGGAACTTCCGACGAACTTGGGGGTGTATGTTGTCTCTGTCCACTTGGATATAC 1142
```

```
Qy 1224 rGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGl 1244
Db 1143 AGGTTAAAGTGTGAACAGACATCGATGAGTGAGCCACTGCTTCCCTCAACATGG 1202
Qy 1244 yValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGl 1264
Db 1203 AGTTTGTAAAGACCTAGTTGGGAAATTCATTGTGTGAGTGCCCATCAGGTTACACAGGTCA 1262
Qy 1264 nArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyLysCy 1284
Db 1263 GCGGTGTGAAGAAATATAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1322
Qy 1284 sValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301
Db 1323 TGTGTATGTTGTGCTGGCTATCTTGCATCATGTGTGAAGGATTTGTAGGT 1374
```

RESULT 13
ABK35663
ID ABK35663 standard; cDNA; 1408 BP.
XX
AC ABK35663;
XX
DT 08-MAY-2002 (first entry)
XX
DE cDNA sequence #54 encoding novel human secreted protein.
XX
KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW immune deficiency disorder; blood disorder; inflammatory disorder;
KW infectious disorder; allergic condition; neurodegenerative disorder;
KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200177289-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010232.
XX
PR 06-APR-2000; 2000US-0195605P.
XX
(GENY) GENETICS INST INC.
XX
PA Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;
PI Clark HP, Pechtel K, Howes SH, Resnick RJ, Gullukota K, Graham JR;
XX
DR WPI; 2002-179322/23.
XX
XX Six hundred and twenty three polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for treating
XX immune deficiencies and disorders such as autoimmune disorders.
XX
XX Claim 1; Page 101; 393pp; English.
XX
XX The present invention relates to the isolation of novel cDNA sequences
XX which encode human secreted proteins. The cDNA sequences have been
XX derived from a variety of human tissues. The invention also provides a
XX method for producing proteins from these polynucleotide sequences. The
XX proteins are useful for identifying compounds that modulate their
XX activity and production. The sequences of the invention are useful for
XX treating diseases such as hyperproliferative disorders (e.g. cancer),
XX immune deficiency disorders (e.g. severe combined immunodeficiency
XX (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
XX (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis),
XX infectious disorders (e.g. hepatitis), allergic conditions (e.g. asthma),
XX neurodegenerative disorders (e.g. Alzheimer's disease), liver fibrosis,
XX coagulation disorders (e.g. haemophilia), and tumours. The polynucleotide
XX sequences of the invention are also useful in gene therapy. ABK35610-
XX ABK36232 represent the cDNA sequences of the invention that encode for
XX novel human secreted proteins

SQ Sequence 1408 BP; 456 A; 285 C; 291 G; 376 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.39e-107 Length: 1408
Score: 1977.00 Matches: 362
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 19.35% Indels: 1
DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x ABK35663 (1-1408)

Qy 505 LysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThrIle 524
Db 2 AAGATGTCATCATATCCCCACACCTGTGGCAAGCAGCCAGCAATTTGGAGCAGTC 61
Qy 525 CysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArgCys 544
Db 62 TGCTATGTAGTTGCCGCCAAGGGTTCATTTTATCTGGAGTCAAGAAATGCTGAGATGT 121
Qy 545 ThrThrSerGlyLysTrpAsnValGlyValGlnAlaLysValCysLysAspValGluAla 564
Db 122 ACCACTTCTGGAAATGGAATGTCGGAGTTCCAGGAGCTGTGTGTAAGACGTGGAGGCT 181
Qy 565 ProGlnIleAsnCysProLysAspIleAlaLysThrLeuGluGlnAspSerAla 584
Db 182 CCTCAATCACTGTCTTAGGACATAGAGCTAAGACTCTGGAAACAGCAGATTTCTGCC 241
Qy 585 AsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerValHis 604
Db 242 AATGTTACTTGGCAGATTTCCACACAGCTTAAGACCACTCTGGTGAAGAGTGTGAGTCCAC 301
Qy 605 ValHisProAlaPheThrProProTyrIlePheProIleGlyAspValAlaIleValTyr 624
Db 302 GTTCATCCAGCTTTCCACCCCACTTACCTTTTCCCAATTGGAGATGTTGCTATCGTATAC 361
Qy 625 ThrAlaThrAspLeuSerGlyValGlnAlaSerCysIlePheHisIleLysValIleAsp 644
Db 362 ACGGCACTGACCTATTCGGCAACAGCCAGCTGCAATTTTCCATATCAAGTTATTGAT 421
Qy 645 AlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGluLys 664
Db 422 GCAGAACCCCTGTCTAGATAGCTGTGTGAGATCTCCACTCCCGTCCAGGTTCTCGAGAG 481
Qy 665 ValHisAlaIleAspTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeuVal 684
Db 482 GTACATGCCGCAAGCTGGAGTGGAGCTCAGTTCTCAGACCACTCAGGGCT-GAATTTGGTC 540
Qy 685 IleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGlnTyr 704
Db 541 ATTACCAAGAGTCTATACACAGGAGAGCTTTTCCCTCAAGGGAGACTATAGTACAGTAT 600
Qy 705 ThrAlaThrAspProSerGlyValAsnAsnArgThrCysAspIleHisIleValIleLysGly 724
Db 601 ACAGCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATGTCATAAAGGT 660
Qy 725 SerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAspAsn 744
Db 661 TCTCCCTGTGAATTTCCATTCACACCTGTAAATGGGATTTTATGCACTCCAGATATAT 720
Qy 745 ThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySerThr 764
Db 721 ACTGGAGTCACTGATCATATTAACCTTGTGGAGGGCTATGATTTTCAAGAGGGTCTACT 780
Qy 765 AspLysThrTyrCysAlaLysGluAspGlyValTrpLysProThrThrThrThrIleTrp 784
Db 781 GACAAGTATTATTGTGCTTATGAGATGGGCTCTGGAAACCAACATATACCACTGATGG 840
Qy 785 ProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPheTyr 804
Db 841 CCAGACTGTGCCAAAACACGTTTTCGAACACACGGGTTCAAGGTTCCATTTGAGATGTTCTAC 900
Qy 805 LysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGluThr 824

Db 901 AAGACAGCTGTTGCTGATGACACAGATCTGATGAAGAAGTTTCTGGAAGCATTTGAGAGC 960
Qy 825 ThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArgLeu 844
Db 961 ACCCTGGGAAAAATGGTCCCATCATTTTGTAGTAGTGACAGAGGACATTTGACTGCGAGCTG 1020
Qy 845 GluGluAsnLeuThrLysLysThrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe 864
Db 1021 GAGGAGAACCTGACCAAAAATATTTGCTAGATATATATTTATGACTATATGAATGCTTT 1080
Qy 865 AlaIleGly 867
Db 1081 GCATTTGT 1089
RESULT 14
AAS26857
ID AAS26857 standard; cDNA; 1892 BP.
XX
AC AAS26857;
XX 07-NOV-2001 (first entry)
XX Human cDNA encoding a novel secreted protein, SEQ ID 49.
XX Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
XX cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
XX antibacterial; virucide; fungicide; ophthalmological; vulnerary;
XX secreted protein; rheumatoid arthritis; hyperproliferative disorder;
XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
XX cerebral ischaemia; angiogenesis; nervous system disorder;
XX Alzheimer's disease; infection; ocular disorder; corneal infection;
XX wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
OS Homo sapiens.
XX
XX WO200155441-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001320.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184684P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225577P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.

PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251858P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-476222/51.
 P-PSDB; AAU16952.

Novel polypeptides and polynucleotides useful as diagnostic reagents to
 diagnose diseases or disorders associated with aberrant expression or
 activity of polypeptides, for treating blood clotting disorder,
 hemophilia.

Claim 1; SEQ ID NO 49; 601pp; English.

The invention relates to isolated nucleic acid molecules and their
 encoded secreted proteins. The nucleic acids and proteins are used to
 prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
 in diagnosing a pathological condition or susceptibility to a
 pathological condition. Antibodies to the proteins can also be used in
 alleviating symptoms associated with the disorders and in diagnostic
 immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
 (ELISA). Disorders which are diagnosed or treated include autoimmune
 diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
 neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
 arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
 nervous system disorders e.g. Alzheimer's disease, infections caused by
 bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
 and many other disorders listed in the specification. The polypeptides
 can also be used to aid wound healing and epithelial cell proliferation,
 to prevent skin aging due to sunburn, to maintain organs before
 retransplantation, for supporting cell culture of primary tissues, to
 regenerate tissues and in chemotaxis. The polypeptides can also be used
 as a food additive or preservative to increase or decrease storage
 capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 minerals, cofactors and other nutritional components. The present
 sequence encodes a novel secreted protein of the invention. Note: The

Alignment Scores:

Pred. No.: 3 81e-79 Length: 1892
 Score: 1498.00 Matches: 263
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.62% Mismatches: 0
 Query Match: 14.66% Indels: 0
 DB: 4 Gaps: 0

US-09-977-053-6 (1-1842) x AAS26857 (1-1892)

QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
 DB 2 TGTGTTGATGGTGGCTGCTATCGTGCACATGTGAAAGGATTTGTAGCGCTGCAT 61
 QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu 1323
 DB 62 TGTGAACAGAAAGTCAATGAATCCAGTCAAAACCCATGCTTAAATATGAGTCTGTGAA 131
 QY 1324 AspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
 DB 122 GACCAGTTGGGGATTCATGTCMAATGCCACCTGGATTTTGGTACCCGATGTGA 181
 QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
 DB 182 AGAAGCTGCATGAGTGTCTCAGTCAGCCATGCAGAAATGAGCTACTCTGAAGACGGT 241
 QY 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
 DB 242 GCCAATAGCTTCAGATGCTGTGTGAGCTGGCTTCACAGATCACACTGTGAATGAAC 301
 QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
 DB 302 ATCAATGAATGTCAGTCTAATCATCATAGTAATCAGCCACCTGTGTGGATGAATTAAT 361
 QY 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
 DB 362 TCATACAGTTGTAATGTCAGCCAGGATTTTCAGGCCAAGGTCGAAACAGAACAGTCT 421
 QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyLysLeuGlyTyrValMetLeuAspGly 1443
 DB 422 ACAGGCTTTAACTGATTTTGAAGTTCTGGCATCTATGATATGATGCTAGATGTC 481
 QY 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAspMet 1463
 DB 482 ATGCTCCATCTCTCCATGCTTAACCTGATCTTCTGGATGAATCTCTGACGATG 541
 QY 1464 AsnTyrGlyThrProLeuSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
 DB 542 AACTATGGAACCAATCTCTATGCAATGATTAACCGCAGCGACCAATACCTTGCTCTG 601
 QY 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluValLeuThrAsnCysPro 1503
 DB 602 ACTGATTTAAACGGCTGGGTTCTTTATGTAATGGCAGGAAAGATAACAACTGTCCC 661
 QY 1504 SerValAsnAspGlyArgTrpHisHisAlaLeuThrTrpThrSerAlaAsnGlyLeu 1523
 DB 662 TCGGTGATGATGCGAGATGCGATCATATTGCAATCACTTGGACAGTGGCAATGCAATC 721
 QY 1524 TrpLysValTyrTrpLeuAspGlyLysLeuSerAspGlyAlaGlyLeuSerValGlyLeu 1543
 DB 722 TGGAAAGTCTATATCATGATGGAAATATCTGACGGTGGTCTGCGCTCTCTGTGGTTG 781
 QY 1544 ProLeuProGly 1547
 DB 782 CCCATACCTGGT 793

RESULT 15

AAH08402

ID AAH08402 standard; cDNA; 765 BP.

XX AC AAH08402;

XX DT 26-JUN-2001 (first entry)

XX

DE Human cDNA clone (5'-primer) SEQ ID NO:5237.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX Homo sapiens.
 XX EF1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 99JP-00248036.
 XX 27-AUG-1999; 99JP-00300253.
 XX 11-JAN-2000; 2000JP-00118776.
 XX 02-MAY-2000; 2000JP-00183767.
 XX 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 1; SEQ ID NO 5237; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 765 BP; 202 A; 182 C; 194 G; 183 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.28e-76 Length: 765
 Score: 1447.00 Matches: 251
 Percent Similarity: 98.82% Conservative: 0
 Best Local Similarity: 98.82% Mismatches: 3
 Query Match: 14.16% Indels: 0
 DB: 4 Gaps: 0

US-09-977-053-6 (1-1842) x AAH08402 (1-765)

QY 195 ThrAspGlyTyrSerAsnGlyGlyAspProArgProIleAlaSerLeuArgAspSer 214

DB 2 ACTGATGATATTCCCAATGGGGAGACCCCTAGACCAATTGCGCGTCACTCGAGATTCA 61

QY 215 GlyValGluIlePheThrPheGlyLeuTrpGlnGlyAsnIleArgGluLeuAsnAspMet 234

Db 62 GGAGTGGAGATCTTTCACATTTTGGCATATGGCAAGGACATTCGAGAGCTGAATGACATG 121
 Qy 235 AlaSerThrProGluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAla 254
 Db 122 GCTTCCACCCCAAGGAGGAGCAGCTTACCTGCTACACAGTCTTGAAGAAATTTGAGGCT 181
 Qy 255 LeuAlaArgArgAlaLeuHisGluLeuLeuProSerGlySerPheLeuGluAspMet 274
 Db 182 TTAGCTGCGCGGCAATGATGAAGATCTTACCTCTCGGAGGTTTATTCAGATGATATG 241
 Qy 275 ValHisCysSerTyrLeuLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 294
 Db 242 GTCCACTCTCATATCTTTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
 Qy 295 LysCysGlyThrHisThrGlyHisPheGluGluGluGluGluGluGluGluGluGluGlu 314
 Db 302 AAATGTGGGACACACACAGGCCATTTTGAAGTGCATCTGTGAAAGGGGTATTACGGGAA 361
 Qy 315 GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro 334
 Db 362 GGTCTGCAAGTATGAATGACAGCTTGCCTCGGAGGACATACAAACCTGAAGGCTCACC 421
 Qy 335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGlySer 354
 Db 422 GGAGGAATCAGCAGTGTGATTCATTCATGCTGATGAATAATCACAACCTCCACTGGAAGC 481
 Qy 355 ThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGlu 374
 Db 482 ACATCCCTGGAAGCTGTCTGACAGAGAGGAGGATACAGGGCATCTGGGCGAGACCTGTGAA 541
 Qy 375 LeuValHisCysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCys 394
 Db 542 CTGTGCCACTGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 601
 Qy 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
 Db 602 AACACCACTTCATGACAGCTGTGGGTCCGATGTCAACCTGGATTTGATCTGTGGGA 661
 Qy 415 SerSerIleLeuLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArg 434
 Db 662 AGCAGCATCATCTTATGCTACCCAAATGTTTGTGTCGGTCCGNTTANAGAGCTTACTGCAGA 721
 Qy 435 ValArgThrCysProHisLeuArgGlnProLysHisGlyHis 448
 Db 722 GTAAGAACATGCTCTCATCTTCGGCAGCGCAACATGNCAC 763

RESULT 16

ID ABL14889 standard; cDNA; 10489 BP.
 AC ABL14889;
 DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 39149.
 XX Drosophila melanogaster expressed polynucleotide; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.

XX WO200171042-A2.
 XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.
 DR P-PSDB; ABB70786.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 39149; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 10489 BP; 2233 A; 3116 C; 3060 G; 2080 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 9,93e-70 Length: 10489
 Score: 1354.00 Matches: 514
 Percent Similarity: 32.61% Conservative: 250
 Best Local Similarity: 21.94% Mismatches: 705
 Query Match: 13.25% Indels: 879
 DB: 4 Gaps: 87
 US-09-977-053-6 (1-1842) x ABL14889 (1-10489)
 Qy 6 AlaPheCysTyrTrpGlyLeuAlaLeuValSerGlyTyrPheGlnGlnMetSer 25
 Db 1417 GCACAATGCTGATGTCGCGCACATTCGCGC-----1452
 Qy 26 ProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaProGlySer 45
 Db 1453 -----CGATTGC--CGAGCTGCAACCCGGA-----1477
 Qy 46 IleProAlaProProAlaProGlyAspGluAlaAlaGlySerArgValGlu-----62
 Db 1478 ATTGGAGCTTGGCTTCCAGTCGCGACCGCTTTTCGGTACGCTGCTCAGCTTTACATGT 1537
 Qy 63 ArgLeuGlyGlnAlaPheArgArg---ArgValArgLeuLeuArgGluLeuSerGluArg 81
 Db 1538 CCCATTGGACAGGAGTTTGCACCGCGCAAGACGCGACTGGTTACCGGAATG-TCTGCGCGG 1596
 Qy 82 LeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSer 101
 Db 1597 TGGCAACTGGAGTGTCTCTCATATACC-----CAAGTGTGAGG-AGGTCT 1640
 Qy 102 GluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArg 121
 Db 1641 ACTGCGGTCTGTGCCACAATCGACACGCTTTCATTCGCTGCTCTCGAAGCTAACCT 1700
 Qy 122 ValAlaIleValThrPheSerSerLysSerValValProArgValAspTyrIleSer 141
 Db 1701 ATCGCGGTATAGCAATGTACCACTGCTACGCGCGCTTTTCGCTCGCGTGCCTCGGA 1760
 Qy 142 ThrArgArgAla-----ArgGlnHisLysCys---Ala 151
 Db 1761 TCGAAGAGATC-TCTGTCTCGGATGGCCGTTGGAGCGAGACCCCACTGATGGCC 1819
 Qy 152 LeuLeuLeuGlnGluIleProAlaIleSerTyr-----ArgGlyGly 165
 Db 1820 TCCAGTGGCAGCGCTGCGGAAGTGGCACACGCAACGTCACCTGCTGTAATGAGGT 1879
 Qy 166 GlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAlaArgGlu 185
 Db 1880 GGT-----CGCAGC 1888

186 AsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAspProArg 205
1889 TAGCGACCATTTGTCAGTATGATGAGTGTGAGCCGGGTACGAGCGCATGCGCATCCC--- 1945
206 ProIleAlaLeuSerLeuArgAspSerGlyValIlePheThrPheGlyIleTyrGln 225
1946 -----GTGCTGACCTGTATGTGCGAAGCGCATCTGGAGT 1978
226 GlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCysTyrLeu 245
1979 GGTGATGTATCAAGATGCACG-----CGCAAGCGGTGCTTC--- 2014
246 LeuHisSerPheGluGluPheGluAlaLeuAla-----Arg 257
2015 -----GAAATCCGACCATTTGCCAAGCGCTTTGTGTGGACTCGACCGGA 2059
258 ArgAlaLeuHisGluAsp-----LeuProSerGlySerPheIleGln 271
2060 GCCTACTCTTCGGCGATGAGGCGAGGTGCTGCTTCAAGGGCTACAACTGATCGGC 2119
272 AspAspMetValHisCysSerTyr-----LeuCysAspGlu 283
2120 AGCAACATCATGCGCTGACGAGCGAGGCCAGAAAGTTCGAGCAGCGCGCGAGTCCGAGGAC 2179
284 GlyLysAspCys-----CysAspArgMetGlySerCysLysCysGlyThrHis 299
2180 ATCAACAGGTGAGCTCTCTCGAGTGGAC-----CTAACACACCGAGTCCAGAACACG 2236
300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrProly----- 313
2237 AACGGCTCTCTCACTCCAGTGCAGGACGCGGATTCAC--GGCTACCACCGAGTGTGGGCC 2295
313 ----- 313
2296 CGTCGTGATTTGGGCTTGGGTAAATGAGGCGATACCGGATGACAGCATCACCACTCGGT 2355
314 -----LysGlyLeuGlnTyrGluCysThrAla----- 322
2356 CAGTGAGCCGGGTACAGCAGGAGCAGCTGCGCTTGAACACGATGGCTGGTGGCGGTGG 2415
323 -----CysProSerGlyThrTyrLysPProGluGlySer----- 333
2416 CTCTTCGGAGCTGGTGTGCCAAGTCTGATCTACTCGACCTGAGCGCACCCACCATTTCTGCG 2475
333 ----- 333
2476 TGGCTTTCCGACCATGTCGTGCGAGCGTCCCGATGGCAATGTGGCTTCAGCTCGCGCGT 2535
333 ----- 333
2536 GCGTCTGCGAGTACACCAACGATCTGACGGATGTGTTCAAGGATTATGCCAATCCCGACGG 2595
333 ----- 333
2596 CACTGCGGTGCAATTCCGCATCTCGGAGCCAGCTCTCCATCTTAAACCTGCGCCCTGCC 2655
334 -----ProGly----- 335
2656 CATCGAAGCTCGTATATTGCTTCGCGATCCAGGACTAGTGGGTGGCGCCCTGTCTGGG 2715
336 ---Gly-IleSerSerCys-----IleProCysProAspGluAsnHisThrSerProPr 352
2716 CATGGAGCTGATGGGTGACGCGCTTGGATTGCTGGACATCAACGAGTGCAGCAAGAA 2775
352 oglySerThrSerProGluAspCysVal-----CysAr 363
2776 CAATGCGCGTGTGACCAAGAGTGCATCAACTCACCGGGGGATTTGCTGTGGCTGCGCAA 2835
363 gGluGlyTyrArg----- 368
2836 CACTGGCTACCACTGTATACCTTCCAAACGCGAGCGGTGCTATCACATCGAAGCGCTCGA 2895
368 aserGlyGln-----ThrCysGluLeuValHisCys 378

2896 ATCCGGCAACAGTGTGTCACACCTATCAGCGCAACAGACCTGTGTCTCTCATGTG 2955
378 sProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPh 398
2956 TCCCGAACTGGAGCGCGCCGAGAATGTCACCTCTCAGCGCAGCAAGAACATCATCTT 3015
398 eAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleI 418
3016 TGGCGATGTGGTGGCTTCCAGTGGCACTTTGGCTACATCATGAGCGGACGCTCGGGCG 3075
418 eLeuCysLeuProAsnGlyLeuTyrPheSerGlySerGlySerTyrCysArgValArgThrCy 438
3076 CCTGTGCTCTCCAGCGGTCACTGGAACGCCAGGTACCGAGTGCATATTATGCAATG 3135
438 sProHisLeu----- 441
3136 CGTTCCCTGCCCGATGACAAAGTTGAGGGGTCTGACTGTGGGCCCGCCCGATCCCGAATC 3195
441 ----- 441
3196 CGTTCTAGTGCCTTCCGTGACAAATGTGACCAATTACGTGGGATCGCGGAGCGCCAACT 3255
441 ----- 441
3256 GAGAGCCACCGCTTCTCTGTTTCCGCGAGTGCCTGTAAGATCCCAAGCCCGGTCTGCC 3315
442 -----ArgGlnProLysHisGlyHisIleSerCys----- 451
3316 CGATTACTGCTATCCGAATGACGCGCTTGTGCCCGAGTGATGTCTACTACCCCAT 3375
452 -----SerThrArgGluMetLeuTyrLysTh 460
3376 GCCAACCCCGCGCGCAGAAATACGACAGTTTGTGGACACTGCG-----TATCAGAG 3426
460 rThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly-----Se 475
3427 CAGCTTCTTCTTGGCTGCCAGAACACCTTTAAGTTGGCTGACAGACGCGGTCTGCACGA 3486
475 rAspLysLeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValG 495
3487 CAATGTGTTGTGTGGAGCCGATGTATCTGGGACTTTGGAGATCTTCTGCTGTGAGG 3546
495 uArgHisCysSerThrPheGlnMetProLysAsp----- 506
3547 ACCTGTGTGCGAGATCCGGAGACCGCGAGATGGTGGCGAGATTGCACCGCAGTATGA 3606
507 -----ValIleIleSerProHi 512
3607 GCAGAGCTCGGAGGTGTACTTCCGCTGCAATCGTCTGCTGTACATCTCTGATCAATCCGCG 3666
512 sAsn----- 513
3667 ACCCATTTACATGATACGAGCGCAGAGTGCAGGTGCAGGTTCATCAAGCTTTGGGATTAAGTTC 3726
513 ----- 513
3727 CGSCAGGATTCGGATTCCGATTCCGCAATCAATGCCACTTCGAGCGACCCCAATTACGAGGCCAA 3786
514 -----CysGlyLysGln----- 517
3787 GAACATCCGCTCTCACTCGGCCACTGGCTGTGTGGCAAGCAGGAGGCGCTTCACCTATGT 3846
517 ----- 517
3847 GAGCGTGGATCTGGGTGCAGATCTATCGAGTCAAGCGGATTCGTGTGAGGGGTGTGGTTAC 3906
517 ----- 517
3907 CAACGACATTTGTGGGCGAGGCCCGACGGAGATTCGGTTCTTCTACAAACAGCTGAGAGCGA 3966
518 -----ProAlaLysPheG 522
:::|

Db 3967 GAACTACGTGGTGTACTTCCCAATTTCATCTGACCATGCGAGATCCAGGCAACTACGG 4026
Qy 522 yThrIleCysTyrrValser-----CysArgGlnGlyPheIleLeuSerGlyVa 538
Db 4027 CGAGCTGGCCATGATACGCTGCCCAAGTTCTGTGAGCGCTCGCTTTGTGATCTTGGAA 4086
Qy 538 lLys-----GluMetLeuArgCysThrSerGl 548
Db 4087 AGTGAGCTACATGGACAACGCTGTCTGAGTTTCAGTTGATGGCTGCGAGGACCGAA 4146
Qy 548 yLys-----TrpAsnValGlyValGlnAlaValCysLysAspValGl 563
Db 4147 ACAGGACCACTCTCTCGGCTAGACTACGGCTAC-----TCCCGCTGCTGGACAACGA 4200
Qy 563 ualAProGlnIle---AsnCysProLysAsp---IleGluAlaLysThrLeuGluGlnGl 581
Db 4201 ACCACCCATCTTCCAAATCTGCCCGCAGCAACCAATTTGTGGCCACCGCATGAGAA 4260
Qy 581 nAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysVa 601
Db 4261 AGGAGTACTACCGTTAACTTACCGAACCCACGCGGCTGGACAACTCGGATCGATTGC 4320
Qy 601 lSerValHisValHisProAla-----PheThrProProTyrrLeuPheProIleGlyAs 619
Db 4321 CGCCTGGAGATCAAGCCACAGAACTTCGCGACACCCAGCTACATTTTCAAG-----GA 4374
Qy 619 pValAlaIleValTyrrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHi 639
Db 4375 TACGGTTTAAAGTACGTGGCGCTTTGACTACGATGGCAATGTGGCCATCTCGGAGATCAA 4434
Qy 639 sIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProProProVa 659
Db 4435 CATCACGGTGGCCGATGTAAACACACCACTGCTGCAG---TGC-----CCCCAGAG 4482
Qy 659 lGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSe 679
Db 4483 CTATGTGATTGAGCTAGTGGATGGCAGCAGCAGCTACACTGTGAACATTCAACGATACCCG 4542
Qy 679 rGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu-----Ph 695
Db 4543 GAAGAGG-----ATCAAGACCTCCGACGACACAGAGATGTGAGTTGCATTCAG 4593
Qy 695 eProGlnGlyGluThrIle-----ValGlnTyrrThrAlaThrAs 708
Db 4594 CCCGAGAGTCCCAATCAAGATCGGAACCTTCAGAACTGACCGTTCAGCGCAACCGA 4653
Qy 708 pProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySerProCys-- 727
Db 4654 TAAGTACAACAACCGCGCGCTGCCACTTCAGGTCTCTGTGAAGGCTTCACCTGCGT 4713
Qy 728 -----GluIleProPheThrProValAsnGlyAspPheIleCysThrProAspAsnTh 745
Db 4714 GGACTGGGAGCTCCAG-----CGCCCGCGAATGTGTGCCATCAATTGCCCTGGTGATCG 4770
Qy 745 rGlyValAsnCysThrLeuThrCysLeuGluGlyTyrrAspPheThrGluGlySerThrAs 765
Db 4771 TGGTATCGAATGCAATGTCAGATCGAACCCAGGATTCGTTTACCGACGCGGACCACT 4830
Qy 765 pLysTyrrTyrrCysAlaTyrrGluAspGlyValTrpLysProThrTyrrThrGluTrpPr 785
Db 4831 GAAGACCTTCTCTCGAGACATCATCGTCTGTGGCTGCCAGCTCCGCTGGTG-----CC 4884
Qy 785 oAspCysAlaLysLysArgPheAlaAsnHisGlyPheLys---SerPheGluMetPheTy 804
Db 4885 CGACTTCGTGTCGGAACACGAGCAGCGCCCTTACACCGTACCGCCCTCCATTTACCTA 4944
Qy 804 rLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGluTh 824
Db 4945 CCGGCCAATGAGCAGTGGCCCAATCTCTGTCTGGTTCAGTACCGAGAGGTCTGGCACA 5004
Qy 824 rThrLeuGly-----LysMetValProSerPheCysSerAsp----- 836
Db 5005 GCACATATGGCGGATCAACACAGTTGCTCTCGACGGCTGCTCGCGCGTGAATGTCAACAT 5064

Qy 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTy 852
Db 5065 GAATGTGACCTTTTGAAGTCTGTGCCCATGCTGCTGGAGGAAATGTGTCAAGATG-- 5122
Qy 852 rCysLeuGluTyrrAsnTyrrAspTyrrGluAsnGlyPheAlaIleGlyProGlyTyrrGl 872
Db 5122 ----- 5122
Qy 872 yAlaAlaAsnArgLeuAspTyrrSerTyrrAspAspPheLeuAspThrValGlnGluThrAl 892
Db 5123 -----GACTTCATCTCTCCAT-- 5140
Qy 892 aThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTy 912
Db 5141 -----CTGCCCGCTGTGGTTCAGCGCAGCTGTACGACTGTGGCGGCTC 5184
Qy 912 rLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspGlu---ArgAs 931
Db 5185 CACGCTGAACCTGATCTTTGATCTG-----AGTGTACCTATGCCAGTGCCTGATCGA 5238
Qy 931 nAspThrLeuGluTyrrGluAsnGlnArgLeuLeuGlnThrLeuGluThrIleThrAs 951
Db 5239 TGACCTTTGAACATTCGCAACATCGTAACCAAGTGTCTCGCTACGCGCCCTCAAGTC 5298
Qy 951 nLysLeuLysArgThrLeuAsnLysAspProMetTyrrSerPheGlnLeuAlaSerGluIl 971
Db 5299 GCAATCTCGCGAGATTTAACTGCAAT-----GTGGCGCAGGT 5337
Qy 971 eLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerProPheCysArgProGl 991
Db 5338 ACTGAACATGGACACCGCATGTCCG----- 5365
Qy 991 ySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrrTyrrAsnLeuGl 1011
Db 5366 -----CGTTGCTGCACCTGTCCCGCGAAGCTATGTGTGACAGGG 5406
Qy 1011 uHisPheThrCysGluSerCysArgIleGlySerTyrrGlnAspGluGluGlnLeuGl 1031
Db 5407 TCAGAACAGCTTGACCTACTTCCCGAGGGGCTTACTACAGAACCGTGACCGCGGGAAC 5466
Qy 1031 uCysLysLeuCysProSerGlyMetTyrrThrIleHisSerArgAsnIleSerAs 1051
Db 5467 CTGCTGCGCTGCCCGCGCGGAACCTACCAAGAGGAGGAGGCGCACCAAGTCGACGCGGA 5526
Qy 1051 pCysLysAlaGlnCysLysGlnGlyThrTyrrSerTyrrSerGlyLeuGluThrCysGluSe 1071
Db 5527 CTGCATTCCTGCTGCGGTTATGGCACCTTACTCACCACCGGACTGGTCCGCTCGGA 5586
Qy 1071 rCysProLeuGlyThrTyrr-----GlnProLysPheGlySerArgSerCysLeuSe 1088
Db 5587 GTGTCCGCTAACTCATTCTACCTGCGCAACCAACCGGTGATTCAGGATTCGACGCG 5646
Qy 1088 rCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGlyVa 1108
Db 5647 CTGTCCGCGCACAGACTTCACTACAGCGCGCTGCTCGAACAAAGGATCTGTGTGCGCG 5706
Qy 1108 lProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisProCysProAr 1128
Db 5707 CAAGTGTGCGCGGGAACCTACTTCCCGCAGGACTGGACCTGCTGCCCTGCGCCACT 5766
Qy 1128 gAspTyrrTyrrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPheTyrrGlyTh 1148
Db 5767 GCATCATTCACAGGAGCGCGGGTGGCAGAGCTGCAACGAGTGTCCGAGTAACATGAG 5826
Qy 1148 rThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSerThrPheSe 1168
Db 5827 AACCGATTCCACCGCTTCAAGGAGCAGCGAACAGTGC----- 5863
Qy 1168 rAlaAlaGluGluSerValProProAlaSerLeuGlyHisIleLysLysArgHisGl 1188
Db 5864 -----AAGCGGTGGTA----- 5975

1188 uileSerSerGlnValPheHisGluCysPheAsnProCysHisAsnSerGlyThrCy 1208
5876 -----TGTGGTGAAGGTGCTTGCAGCAGCGCGACTGTG 5910
1208 sGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCy 1228
5911 TGTGCCCATGGCCATGACATCCAGTGTCTGTCCGCGCGGATCTCTGAGCGCTGCGTG 5970
1228 sGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGlyValCysLysAs 1248
5971 CGAACAGGATCGACAGTGGCTCCCGACCTCTGTACAAATGGTGTGATGCAAGGA 6030
1248 pLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGlnArgCysGluG 1268
6031 TCTGCCGAGGGCTATCGCTGTGAGTGGCGGCTGGATCTCGGGCATCAATTCGAGGA 6090
1268 uAsnIleAsnGluCysSerSerProCysLeuAsnLysGlyIleCysValAsp----- 1286
6091 GGAGCCAGTGTACTGTGGCAACGACACTGTCTCCGCGCAGGGCCATGTGCAAGAACGAGCC 6150
1287 -----GlyValAlaGly----- 1290
6151 GGCGTACAGAACTGACCTGTCTGTGCGCGAGTGGCTTACACCGCGGATCAGTCCGACGT 6210
1290 ----- 1290
6211 GACCATCGATCCGTGCACGGCGAATGCATCCGTGGCGAAACGGAGCCAGCTCCAGGC 6270
1291 -----TyrArgCysThrCysValLysGlyPheValGlyLeuHisCysG 1305
6271 CTTGGAGCAGGTGCTGTACAGTGCAGTGTGTCCCGGATGGGAGGCGCATCTCTGA 6330
1305 uThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGluAspG 1325
6331 GCAGAAATATCAATGATGTTCCGAGATCCCTGCTGTGGCGCCCAACTGCACAGATCT 6390
1325 nValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGlyLysAs 1345
6391 GGTCAATGACTTCCAGTGGCGCTGTCCGCCAGGATTTACGGCGAAGCGATCGAGCAAA 6450
1345 nValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGlyAlaAs 1365
6451 GATCGATCTCTGCTATCGCAACCATGCAAGCATGGC-----ACCTGCGTGGATCTCTGTT 6507
1365 nSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAs 1385
6508 CGATCAGCAGTGTGTTGGCCATCCGCGCTGACGGGATCCGCTCGCACATCAACATCGA 6567
1385 nGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTy 1405
6568 CGACTCGAGAACCGACCTCGGCCAATGAGGNACTTGGTGCACCTGCTGCGAGGCTA 6627
1405 rSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu-ThrGluGlnSerThrG 1425
6628 TAGCTGCAACTGTGAACCCGCTACACGGCGCAGAAATTCGACGACACCATCGACGACTG 6687
1425 lyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGlyMetL 1445
6688 CG----- 6689
1445 euProSerLeuHisAlaLeuThrCys-----ThrPheTrpMetLysSerSerA 1461
6690 --CCTCGAATCCCTCCAGCAGCGCGCCACTGTGTGACCACTGGTGCACCTGCTGAGCTG 6747
1461 spAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrL 1481
6748 CAAATCGCGCCCTGGCTACGCTGGGTCTCTCTCG----- 6782
1481 euLeuLeuThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyArgGluLysIleThrA 1501
6783 -----AGCCGAGATCGACGA----- 6798
1501 snCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAlaA 1521

6799 GTGTCGAGCGACCCCTCGCAATCCGGTGGCGACGG-----AGCGCTG 6840
1521 snGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer- 1540
6841 CCTCGATCTGGACAAACAAT-----TCGAGTGGCTGTGTCGGGACGGATT 6885
1541 ValGlyLeuProIleProGlyGly-----GlyAlaLeuVal-Leu----- 1553
6886 CAAGGACCCCTGTGCGCCACGACATCGATCGAGGCGGCGAGCCGCTGTCTGAACAA 5945
1554 -GlyGlnGluGlnAspLysLysGlyGlu-----GlyPheSerProAlaGluSerPh 1570
6946 CGGCATCTGTCCGATCGCGTGGCTTGGAGTGGCGCTCGCGAGCCAGGA----- 6997
1570 eValGlySerIleSerGlnLeuAsnLeuTrpAspTyrValLeuSerProGlnGlnVally 1590
6998 -----TGGAGTGGCATGGCTCGCGACGACGAGTGCAC 7029
1590 sSerLeuAlaThrSerCysProGluLeuSerLysGlyAsnValLeuAlaTrpProAs 1610
7030 CACGTGGAGCTCAGGGCGCTGCCAGAACGATGCCAGCTGCATCGACCTGTTCCAGGA 7089
1610 pPheLeuSerGlyIleValGlyLysValIleAspSerLysSerIlePheCysSerAs 1630
7090 CTAC-----TTCCTCGGTG-- 7102
1630 pCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysPr 1650
7103 -TGTCGCCAGC---GGCACCGATGGCAGACATGCGAGACCGCTCCGGA-- 7147
1650 oGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValG 1670
7148 -----CGCTGCATCGGTGATCTCT----- 7165
1670 nTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGlu----- 1686
7166 ---TGCATGCACGGTGGCAAG-----TGCCAGGACTTTGGCTC 7200
1687 -ArgIleSerCysGlyValProProProLeuGluAsnGlyPheHisSerAlaAspAsph 1706
7201 TGGTCTTAACATGCAATGTCCT-----GCGGAT----- 7228
1706 eTyrAlaGlySerThrValThrTyrGln-----CysAsnAs 1718
7229 -TACTCGGCATTTGGTGTCTAGTACGAGTACGACGATCGCAGGAGCATGCTGTGAGAA 7287
1718 nGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnG 1738
7288 TGGC-----GCCACTGTGTGGACAATGCTGCT-----GG 7317
1738 yValSerProSerCys-----LeuAspValAspG 1748
7318 CTACAGCTGCCAGTGGCCACCTGCTCCACCGTGGCAATTCGGAACAGGACATCGTGA 7377
1748 uCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTy 1768
7378 CTGCAAGGACAACTCT---TGCCCAACCGGCGCCACGCTGCTGGATCTTAACCAACGGCTT 7434
1768 rIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLy 1788
7435 CTACTGTCACTGTC-----CCCTTCAATATACCGGAGACGATTCGCGCAAGGCATCCA 7488
1788 s 1788
7489 A 7489
RESULT 17
ACA57318
ID ACA57318 standard; cDNA; 706 BP.
XX
AC ACA57318;
XX

DT	10-JUN-2003	(first entry)	
DE	Human adipocyte Selected Interacting domain, SID, cDNA #405.		
KW	Human; ss; gene; prey; adipocyte; SID; selected interacting domain;		
KW	anorectic; antidiabetic; protein-protein interaction; diabetes;		
KW	yeast 2-hybrid assay; metabolic disorder; obesity.		
XX	Homo sapiens.		
OS	WO200286122-A2.		
PN	31-OCT-2002.		
PD	14-MAR-2002; 2002WO-EP003768.		
PF	14-MAR-2001; 2001US-0275734P.		
PR	(HYBR-) HYBRIGENICS.		
PA	Legrain P, Daviet L;		
PI	WPI; 2003-103412/09.		
DR	P-PSDB; ABB07074.		
XX	New complex between two interacting proteins in adipocyte cells, useful		
PT	for identifying selected interacting domains that modulate protein		
PT	interactions, or for preventing or treating metabolic disorders such as		
PT	obesity or diabetes.		
XX	Claim 7; Page 244; 382pp; English.		
PS	The invention relates to a complex between two interacting proteins in		
CC	adipocyte cells, given in the specification. The proteins are identified		
CC	by selecting a bait protein from a known adipocyte marker and then		
CC	performing a yeast 2-hybrid selection to isolate prey proteins encoded by		
CC	members of an adipocyte cDNA library. The proteins are designated SID		
CC	(RTM) (selected interacting domains) proteins. Also included are a		
CC	polynucleotide encoding a polypeptide in the adipocyte cells, a		
CC	recombinant host cell expressing at least one of the interacting		
CC	polypeptides of the complex, selecting a modulating compound in acid		
CC	cells, a SID (RTM) polypeptide comprising any of the 738 amino acid		
CC	sequences given in the specification (including its fragment or variant),		
CC	a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences		
CC	given in the specification (including its fragment or variant), a vector		
CC	comprising the SID (RTM) polynucleotide, a recombinant host cell		
CC	comprising the vector, a protein chip comprising the polypeptides and a		
CC	record comprising all or part of the data, listed in the specification.		
CC	The complex, polypeptides, polynucleotides and compounds are useful for		
CC	preventing or treating metabolic disorders such as obesity or diabetes.		
CC	The polynucleotides are useful as probes or primers. The complex is		
CC	particularly useful for identifying selected interacting domains (SID		
CC	(RTM)) for screening drugs that modulate the protein interaction, thus		
CC	exhibiting the therapeutic effect. The present sequence encodes a SID		
CC	(prey) protein of the invention		
XX	Sequence 706 BP; 202 A; 137 C; 173 G; 194 T; 0 U; 0 Other;		
SQ	Alignment Scores:		
	Pred. No.: 2, 11e-70 Length: 706		
	Score: 1341.00 Matches: 234		
	Percent Similarity: 100.00% Conservative: 1		
	Best Local Similarity: 99.57% Mismatches: 0		
	Query Match: 13.12% Indels: 0		
	DB: 7 Gaps: 0		
US-09-977-053-6	(1-1842) x ACA57318 (1-706)		
QY	1242 AenAenGlyValCysGlyAspLeuValGluPheIleCysGluCysProSerGlyTyr 1261		
Db	2 AACATGGAGTTGTAAAGACCTAGTTGGGAATTCATTGTGAGTGGCCATCAGGTAC 61		
QY	1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLys 1281		
Db	62 ACAGGTGAGCGGTGTGAAGAAATATAAATGAGTGTAGCTCCAGTCTCTGTTTAAATAAA 121		
QY	1282 GlyTleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301		
Db	122 GGAATCTGTGTGTGATGGTGTGGCTGCTATCGTTGCACATGTGTGAAAGGATTTGTAGGC 181		
QY	1302 LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321		
Db	182 CTGATTGTGAAACAGAGTCAATGATGCCAGTCAAAACCCATGCTTAAATATGAGTTC 241		
QY	1322 CysGluAspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArg 1341		
Db	242 TGTGAAGACCAAGTTGGGGGATTTCTGTGCAATGCCCACTGGATTTTGGGTACCGA 301		
QY	1342 CysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLys 1361		
Db	302 TGTGAAAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGAGAGTACCTGTGAAA 361		
QY	1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381		
Db	362 GACGGTGCCTAATAGCTTCAGATGCTGTGAGCTGGCTTCACAGGATCACACTGTGAA 421		
QY	1382 LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401		
Db	422 TTGAACATCAATGAATGTCACTTAATCATGTAGAAATCAGGCCACCTGTGTGGATGAA 481		
QY	1402 LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421		
Db	482 TTAATTCATACAGTTGTAAATGTCCAGCCAGGATTTTCAGGCAAAAGGTGTGAAACAGAA 541		
QY	1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441		
Db	542 CAGTCTACAGGCTTTAACTGGATTTTGAAGTTTCTGGCACTATGATGATATGTGATGCTA 601		
QY	1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAsp 1461		
Db	602 GATGGCATGCTCCCATCTCTCCATGCTTAACCTGTACCTTCTGGATGAAATCCTCTGAC 661		
QY	1462 AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGly 1476		
Db	662 GACATGAACACTGGAACACCACTCTCTATGATGAGTTGATTAACGGC 706		
RESULT 18			
ABL14888/c			
ID	ABL14888 standard; cDNA; 13484 BP.		
XX	ABL14888;		
AC	26-MAR-2002 (first entry)		
DT	Drosophila melanogaster expressed polynucleotide SEQ ID NO 39146.		
DE	Drosophila; developmental biology; cell signalling; insecticide;		
KW	pharmaceutical; gene; ss.		
KW	Drosophila melanogaster.		
XX	WO200171042-A2.		
PN	27-SEP-2001.		
PD	23-MAR-2001; 2001WO-US009231.		
XX	23-MAR-2000; 2000US-0191637P.		
PR	11-JUL-2000; 2000US-00614150.		
XX	(PEKE) PE CORP NY.		
PA	Venter JC, Adams M, Li FWD, Myers EW;		
PI	WPI; 2001-656860/75.		
DR	P-PSDB; ABB07075.		

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 39146; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-
CC AB872072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 13484 BP; 2974 A; 3619 C; 3741 G; 3150 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,63e-68 Length: 13484
Score: 1328.00 Matches: 518
Percent Similarity: 32.44% Conservative: 256
Best Local Similarity: 21.71% Mismatches: 718
Query Match: 13.00% Indels: 900
DB: 4 Gaps: 88

US-09-977-053-6 (1-1842) x AB114888 (1-13484)

QY 6 AlaPheCysCysThrGlyLeuAlaLeuVal-----SerGlyTyrAla 19
DB 11068 GCACATGCTGCTGGGTAAGTCTTTATATTCAGATTACATAGTTCAATCATCT 11009

QY 20 ThrPheGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAla 39
DB 11008 TCAATCTTGATCTCTAGATGGTGGCCACATCTCTGGCCGATGTC-CCGAGCTGCAA 10950

QY 40 ProGlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySer 59
DB 10949 CCGCGA-----ATTGGAGCCTTGGCCCTCCAGTCCGACACCGCTTCCGTTACG 10902

QY 60 ArgValGlu-----ArgLeuGlyGlnAlaPheArgArg-----ArgValArgLeuLeu 75
DB 10901 CTGGTCAGCTTTACATGTCCTATGGACAGGAGTTTGCCACCGGCAAGACCGGCTGTT 10842

QY 76 ArgGluLeuSerGluArgLeuGluLeuValPheLeuValAspSerSerSerValGly 95
DB 10841 ACCGAATG-TCTGGCGGTGGCAACTGGAGTGTCTCTACATACCCCAAGTGTGAG--GGT 10785

QY 96 GluValAsnPhe----- 99
DB 10784 GAGTGTGCTTACCTATTTCCATAATTGGATTCCATAATCGATATGCAATCCCTTCCA 10725

QY 100 ----ArgSerGluLeuMetPheValArgGlyLeuLeuSerPheProValValProThr 118
DB 10724 CAGAGCTTACTCGGCTCTGGTGGCCACAAATGCAACGCTTCTCCATTTGGCTCTCGA 10665

QY 119 AlaThrArgValAlaIleValThrPheSerSerSerSerSerSerValValProArgValAsp 138
DB 10664 ACCTACCTATTCGGGTATAGCAATGATACCAGTGTCTAGCGCGGCTTTCCTTCGCTCGG 10605

QY 139 TyrIleSerThrArgAla-----ArgGlnHisLys 149
DB 10604 GTGCTCCGATCGAGAAGATC-TCTGTCTGCTCGGATGGCGTGGGAGCGACAGCCCCAC 10546

QY 150 Cys-----AlaLeuLeuGlnGluIleProAlaIleSerTyr----- 162
DB 10545 TGCATGGCTCCCGAGTGGCGAGCGCTGGCGGAGTGGCAGCCACGCCCAAGCTACCCCTGCTG 10486

QY 163 ArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHis 182
DB 10485 AATGGAGGTGGT----- 10474

QY 183 AlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGly 202
DB 10473 ---CGCAGCTACCGCACCATTTGTCAGTATGAGTGTGAGCGCGCTACGAGCGCAATGCG 10417

QY 203 AspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGly 222
DB 10416 CATCCC-----GTGCTGACCTGTATGTTCGACCGC 10387

QY 223 IleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHis 242
DB 10386 ACCTGAGTGGTGTATGATACCAAGATGCAAG-----CGCAAGCGG 10348

QY 243 CysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAla----- 256
DB 10347 TGTCTC-----GAATTCCTCGACCATTTGCCAACCGCTTTGTGTGTGAGC 10306

QY 257 ----ArgArgAlaLeuHisGluAsp-----LeuProSerGlySer 268
DB 10305 TCGACGGCGAGCTTACTCTTTCGGCGATGAGGCCAGGTGTCAGTGTCTTCAAGGGCTACAA 10246

QY 269 PheIleGlnAspAspMetValHisCysSerTyr-----Leu 280
DB 10245 CTGATCGCGCAGCAACATCATGCGCTGCAGCGAGGCCAGAGTTTCGAGCAGCGCCGACG 10186

QY 281 CysAspGluGlyLysAspCys-----CysAspArgMetGlySerCysLysCys 296
DB 10185 TGCAGGACATCAACGAGTGTGAGTCTCTCGCAGTGGCGAC-----CTAACCCACCCAGTGC 10129

QY 297 GlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly----- 313
DB 10128 CAGNACAGCAACGGCTCTCTCCACTGCCAGTGCAGGCGGAGTTTAC--GGTACCACCGA 10070

QY 313 -----LysGlyLeuGlnTyrGluCysThrAla----- 322
DB 10069 GTGTGGCCCGTCTGCTGATTTGGGCTTGGGTAATGAGGAGCATACCGGATGACGATCAC 10010

QY 314 -----CysProSerGlyThrTyrLysProGluGlySer--- 333
DB 10009 CACTCGGTGAGTGGAGCGCGGCTACACCAAGGAGCAGCTGCGCTTGAAACACGAATGGTG 9950

QY 323 -----CysProSerGlyThrTyrLysProGluGlySer--- 333
DB 9949 GTGGGTGGCTCTCTCGAGCGCTGGTGCACACTGATGATCTCATCGACCTGAGGACCCAC 9890

QY 333 ----- 333
DB 9889 CATTCGTGGTGGCTTCGGCACCATGTCCGTGCAGCGTCCCGATGGCAATGTGGCCTTCAG 9830

QY 333 ----- 333
DB 9829 CTGGCGGTGGCTGTGTCAGTGTACCAACAGATCTGACGAGTGTTCAGAGGATTTAGCCAA 9770

QY 333 ----- 333
DB 9769 TCCGACGGCACTGCGCTCGAATTCGCAATCTCTGGAGCCCAAGCTCTCCATCTTAAACT 9710

QY 334 -----ProGly----- 335
DB 9709 GCGCTGCCCATCGAAGCTCGCTATATTTCGCTTCGCGATCCAGGACTACGTGGGTGGCGCC 9650

QY 336 -----Gly-IleSerSerCys-----IleProCysProAspGluAsnHis 349
DB 9649 CTGTCTGCGCATGAGCTGATGGGCTGATGGCTGATGGCTGATGGCTGATGGCTGATGGCTG 9590

QY 349 rSerProGlySerThrSerProGluAspCysVal----- 361
DB 9589 CAGCAAGAACAAATGGCGGCTGTGACCAAGATGATCAATCACTACCGGGCGGATTTGCTGCTG 9530

QY 362 ----CysArgGluGlyTyrArg----- 367
DB 9529 TGGCTGCACACATGCTGCTACCGAGCTGTACACTCCCAACCGGACCGGCTGGCTATCATACGA 9470

Db 7352 CATTACCTACCGCGCAATGGAGCAGTGCGCCCAATCTGTCTGGGTGAGTACCAAGAGGT 7293
Qy 821 aPheGluThrThrLeuGly-----LysMetValProSerPheCysSerAsp----- 836
Db 7292 GCTGGCACAGCATATGCGGACTCAACAGTTGCTCTCGCAGCGCTGCTCCGCGGTGAA 7233
Qy 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuTh 849
Db 7232 TGTCAACATGAATGACCTTTGTGAAGTCTGTGCCCATGCTGTGCGAGAGATGTGT 7173
Qy 849 rlyLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProG 869
Db 7172 CAAGATG----- 7166
Qy 869 yGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValG 889
Db 7165 -----GACTTCATCTCTCCAT-- 7148
Qy 889 nGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLe 909
Db 7147 -----CTGCCCGCTGTGGGTTCAGCCCGAGCTGTACGACCT 7113
Qy 909 uSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspG 929
Db 7112 GTGGCGCTCCAGCTGAACCTGTATCTTGTATCTG-----AGTGTACCTATGCCAGTGC 7059
Qy 929 u---ArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeuGlnThrLeuGluTh 948
Db 7058 CCGTATCGATGATGACCTTTGAACATTCGCAACATCGGTAAACAGTGTCTCCGCTACGCGC 6999
Qy 948 rIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAl 968
Db 6998 CCTCAAGTCGCAATCTCGGAGATTAACTGCAAT-----GT 6960
Qy 968 aSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerProPheCy 988
Db 6959 GGGCGAGGTACTGAACATGACACACCGAGCTGCGG----- 6923
Qy 988 sArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTy 1008
Db 6922 -----CGTTCCTGCACTGTCCGCGCACTGTCCGCGGAGCACTGTATGT 6891
Qy 1008 rAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGlu 1028
Db 6890 GTGAGAGGTGACAGACGCTGACCTACTCCCGAGGGGTACTACAGAACCGTACCG 6831
Qy 1028 yGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAs 1048
Db 6830 CCAGGGAACCTGCTGCGCTGCGCGCGCGGACCTACACCAAGGAGGAGGCGCACCAATC 6771
Qy 1048 nIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluTh 1068
Db 6770 GCAGCGGAGTGCATTCCTGCTGCGGTTATGGCACTTACTCAACCCAGCACTGGTGC 6711
Qy 1068 rCysGluSerCysProLeuGlyThrTyr-----GlnProLysPhePheSerArgSe 1085
Db 6710 GTGCTCGGAGTGTCGCGTGAACCTATCTACTCGGACACCAACCGGTGGATTCCAGGA 6651
Qy 1085 rCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAl 1105
Db 6650 TTGCGAGGCTGTCCGCGCACAGAGCTTACCTACCAAGCGGCTGCTCGAACAAGATCT 6591
Qy 1105 aCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisPr 1125
Db 6590 GTGTCCGCGCAAGTGTGCGCGGAAAGTACTCCGCCACCGGACTGGCACCTCTGCTGCC 6531
Qy 1125 oCysProArgAspTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPh 1145
Db 6530 GTGCCCCACTGATCATTTACAGGAGCGCGGGTGGCAGAGCTGCACAGTGTCCGAG 6471
Qy 1145 eTyrGlyThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSerSe 1165
Db 6470 TAAATGAGAACCGATTACCCCGCTCCAGGAGCGGACAGTGC----- 6425

Qy 1165 rThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGlyHisIleLysLy 1185
Db 6424 -----AAGCCGTTGTA----- 6413
Qy 1185 sArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisAsnSe 1205
Db 6412 -----TGTGGTGAAGTGTCTTGCACGACG 6387
Qy 1205 rGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGl 1225
Db 6386 CGGACTGTGTGTCCTATGGCCCATGACATCCAGTGTCTTGTCCGCGCGGATTTCTCTGG 6327
Qy 1225 yLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnGlyVa 1245
Db 6326 AGTTCCTCGTGGAAACAGACATGACGAGTGGCTGCCAGCCCTGCTACATGTGTGTCA 6267
Qy 1245 lCysLysAspLeuValGlyLysPheIleCysGluCysProSerGlyTyrThrGlyGlnAr 1265
Db 6266 GTGCAAGGATCTGCCGCGAGGCTATCGCTGTGAGTGCCTGCTGATCTCGGGCATCAA 6207
Qy 1265 gCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIleCysVa 1285
Db 6206 TTGCCAGGAGGAGGCCAGTGTGACTGTGGCACACACCTGTCCGCGCCAGGGCCATGTGCA 6147
Qy 1285 lAsp-----GlyValAlaGly----- 1290
Db 6146 GAACGAGCCGGGCTACAGAACGTGACCTGTCTGTGCCGAGTGGCTACACCGCGGATCA 6087
Qy 1290 ----- 1290
Db 6086 GTCCGAGCTGACCATCGATCGTCCGCGAGCGAATGGCAATCCGTCGGAACGCGACG 6027
Qy 1291 -----TyrArgCysThrCysValLysGlyPheValGlyLe 1302
Db 6026 CTGCCAGGCTTGGAGCAGGCTGCTACAAAGTGCAGTGTGTGCCGCGATGGGGGCGCAT 5967
Qy 1302 uHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCy 1322
Db 5966 CCACTGTGACGAGAAATCAATGACTGTTCGGAGATCCCTGCTCTGTGGCGGCAACTG 5907
Qy 1322 sGluAspGlnValGlyLysPheLeuCysLysCysProProGlyPheLeuGlyThrArgCy 1342
Db 5906 CACAGATCTGTGTCATGACTTCCAGTGCCTGTGCCGCGAGGATTTACGGGCAAGCGATG 5847
Qy 1342 sGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAs 1362
Db 5846 CGAGCAAAAGATCGATCTCTGCTTATCGGAACCATGCAAGCATGGC---ACCTCGTGGTA 5790
Qy 1362 pGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLe 1382
Db 5789 TCTCTGTTCGATCAGAGTGTGTGTTCCTATCCGCGCTGGACGGATCCGCTCGGACAT 5730
Qy 1382 uAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLe 1402
Db 5729 CAACATCGAGCTCGGAGAACCGACCTCGGCCAATGAGGAGAACCTGGTGGACCTGGT 5670
Qy 1402 uAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu-ThrGluG 1422
Db 5669 CGAGCGGTATAGCTGCAACTGTGAACCCGCTACACCGCGCAAGAATTTGCCAGCACCAT 5610
Qy 1422 lnsrThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuA 1442
Db 5609 CGAGCACTGG----- 5599
Qy 1442 spGlyMetLeuProSerLeuHisAlaLeuThrCys-----ThrPheTrpMetL 1458
Db 5598 -----CCTCGAATCCCTGCCAGCACGCGCCACCTGTGTGGACCACTGGATGG 5550
Qy 1458 ySerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAsnGlySerA 1478
Db 5549 CTTTCAGCTGCAATATCGCGCTGCTGCTACGTGGGTCTCTCTCTCTGCG----- 5506

QY 1478 spAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluL 1498
 Db 5505 -----AGCCGA 5499
 QY 1498 ysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTpt 1518
 Db 5498 GATCGACGAGTCTCGAGGACCCCTGCAATCCGTGGGCACGG-----5455
 QY 1518 hrSerAlaAsnGlyIleTyrIleAspGlyLysLeuSerAspGlyGlyAlaG 1538
 Db 5454 --AGCGCTCGCTCGATCGACACAAAT--TCGAGTGGCTGTGGCG 5412
 QY 1538 lyLeuSer-ValGlyLeuProIleProGlyGly-----GlyAlaLeuVal 1552
 Db 5411 GGACGATTCAGGACCCCTGTGCGCCAGCATCATGATGCTGCGAGCGACCCGTG 5352
 QY 1553 -Leu-----GlyGlnGluGlnAspLysGlyGlu-----GlyPheSerProAl 1567
 Db 5351 TCTGAACAACGCGATCTGCGGATCGCTCGTGGCTTTGAGTGGCGCTGCGAGCCAGG 5292
 QY 1567 aGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAspTyrValLeuSerProGl 1587
 Db 5291 A-----TCGAGTGGCATGCGCTGCGAGCA 5268
 QY 1587 nGlnValLysSerLeuAlaThrSerCysProGluLeuSerLysGlyAsnValLeuAl 1607
 Db 5267 GGAGGTGACCATCGGACGCTGCGGCGCGGCGGACGATGCGCAGCTGCAGCT 5208
 QY 1607 aTyrProAspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysIlePhe 1627
 Db 5207 GTTCAGGACTAC-----TT 5193
 QY 1627 eCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAs 1647
 Db 5192 CTGCGTG--TGTCCAGC--GGCAGCATGGCAAGAACTGCGAGACCGCTCCGCA-- 5141
 QY 1647 pleuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAs 1667
 Db 5140 -----CGTCGATCGGTGA 5127
 QY 1667 nProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGlu-- 1686
 Db 5126 TCCT-----TGCATGACGCTGGCAAG-----TGCAGGA 5097
 QY 1687 -----ArgIleSerCysGlyValProProLeuGluAenGlyPheHisSerAl 1703
 Db 5096 CTTTGGCTCTGCTCTTAATGCACTGGCCCT-----GC 5064
 QY 1703 aAspAspPheTyrAlaGlySerThrValThrTyrGln-----1715
 Db 5063 GGAT-----TACTCGGGGCTTGGGTGTGTCAGTACGAGTACGAGCGATGATGT 5010
 QY 1716 -CysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySe 1735
 Db 5009 CTGTCAAGATGCG-----GCCACTTGTGTGCAATGGTGC 4974
 QY 1735 rTrpAsnGlyValSerProSerCys-----LeuAs 1745
 Db 4973 T-----GGCTACAGCTGCGAGTCCACCTGGCTTCCACCGGTGCAATTCGAAACAGGA 4920
 QY 1745 pValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAs 1765
 Db 4919 CATCGTGAGTGCAGGACAACTCT--TGCCACCGGGCGCGCCAGTGGTGGATCTTAC 4863
 QY 1765 pGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGl 1785
 Db 4862 CAACGGCTTCTACTGTGAGTGC-----CCCTTCATATGACCGGAGCAGATTGCGCAA 4809
 QY 1785 uProIleLys 1788
 Db 4808 GGCCATCCAA 4799

RESULT 19

AAH07752
 ID AAH07752 standard; cDNA; 801 BP.
 AC AAH07752;
 DT 26-JUN-2001 (first entry)
 DE Human cDNA clone (5'-primer) SEQ ID NO:4587.
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 OS Homo sapiens.
 XX EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-00116126.
 XX 29-JUL-1999; 93JP-00248036.
 XX 17-AUG-1999; 95JP-00300253.
 XX 11-JAN-2000; 2000JP-00118776.
 XX 02-MAY-2000; 2000JP-00183767.
 XX 09-JUN-2000; 2000JP-00241899.
 XX (HELL-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 XX Claim 1; SEQ ID NO 4587; 2537pp + Sequence Listing; English.
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX Sequence 801 BP; 217 A; 198 C; 194 G; 184 T; 0 U; 8 Other;
 SQ
 Alignment Scores:
 Pred. No.: 3,77e-68 Length: 801
 Score: 1304.00 Matches: 247
 Percent Similarity: 91.18% Conservative: 1
 Best Local Similarity: 90.81% Mismatches: 16
 Query Match: 12.76% Indels: 8
 DB: 4 Gaps: 0
 US-09-977-053-6 (1-1842) x AAH07752 (1-801)

Qy 1484 ThrAspTyrRhenGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
Db |||||
Qy 603 ACTGATTATACCGCTGGGTTCTTTATGTGTAATGCAGGGAAGATAACAACACTGTCCC 662
Db |||||
Qy 1504 SerValAsnAspGlyArg 1509
Db |||||
663 TCGGTGAATGATGCACGA 680

RESULT 21
AAS26930
ID AAS26930 standard; cDNA; 680 BP.
AC AAS26930;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, SEQ ID 122.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective;
KW antibacterial; virucide; fungicide; opthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155441-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001320.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.

PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249298P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476222/51.

P-PSDB; AAU17025.

Novel polypeptides and polynucleotides useful as diagnostic reagents to diagnose diseases or disorders associated with aberrant expression or activity of polypeptides, for treating blood clotting disorder, hemophilia.

Claim 1; SEQ ID NO 122; 601pp; English.

The invention relates to isolated nucleic acid molecules and their encoded secreted proteins. The nucleic acids and proteins are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used in diagnosing a pathological condition or susceptibility to a pathological condition. Antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection, and many other disorders listed in the specification. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, CC minerals, cofactors and other nutritional components. The present CC sequence encodes a novel secreted protein of the invention. Note: The

Alignment Scores:
 Pred. No.: 1,06e-66 Length: 680
 Score: 1278.00 Matches: 224
 Percent Similarity: 99.56% Conservative: 1
 Best Local Similarity: 99.12% Mismatches: 1
 Query Match: 12.51% Indels: 0
 DB: 4 Gaps: 0

US-09-977-053-6 (1-1842) x AAS26930 (1-680)

QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValIleGlyPheValGlyLeuHis 1303
 DB 3 TGTGTTGATGTTGGCTGCTATGTTGCACATGTTGAAAGGATTTGTAGCCTGCAT 62
 QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
 DB 63 TGTGAAACAGAGTCAATGAATGCCAGTCAGTCAAAACCCATGCTTAATAATACAGTCTGTGAA 122
 QY 1324 AspGlnValGlyGlyPheLeuCysIleCysProGlyPheLeuGlyThrArgCysGly 1343
 DB 123 GACCAAGTTGGGGATTTCATGTGCAATGCCACCTGGATTTTGGTACCCGATGTGA 182
 QY 1344 LysAsnValAspGluCysLeuSerGlnProCysIleAsnGlyAlaThrCysIleAspGly 1363
 DB 183 AAGAACGTGCGATGAGTCTCAGTCAGGCATGCCAATAATGGAGTACCTGTTAAAGACGGT 242
 QY 1364 AlaAsnSerPheAtqCysLeuCysAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
 DB 243 GCCAATAGCTTCAGATGCTGTGTGCGAGCTGCTTCACAGGATCACCTGTGAATTGAAC 302
 QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
 DB 303 ATCATGATGATGTCAGTCTAATCCATGTAGAAATCAGGCCACCTGTGTGATGAATTAAT 362
 QY 1404 SerTyrSerCysIleCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
 DB 363 TCATACAGTTGTAAATGTGTCAGCCAGGATTTTCAGGCGMAAAGGTGTGAAACAGAACAGTCT 422
 QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
 DB 423 ACAGGCTTTAACTCGATTTTGAAGTTCTGCGATCTATGGATATGTCATGCTAGATGGC 482
 QY 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetIleSerSerAspMet 1463
 DB 483 ATGCTCCCATCTCTCCATGCTCTAACCTGTACCTCTCGATGAATCTCTGACGACATG 542
 QY 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
 DB 543 AACTATGGAACACCAATCTCTATGCAAGTTGATAACGGCAGCGACAAATACCTTGCTCTCG 602
 QY 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyValGluValIleThrAsnCysPro 1503
 DB 603 ACTGATTATTAACGGCTGGGTCTTTATGGAATGGCAGGGAAGATAACAACTGTCTCC 662
 QY 1504 SerValAsnAspGlyArg 1509
 DB 663 TCGGTGAATGATGGCAGA 680
 RESULT 22
 ABA06507
 ID ABA06507 standard; cDNA; 680 BP.
 AC ABA06507;
 XX
 DT 10-JAN-2002 (first entry)
 XX
 DE Human cDNA SEQ ID NO: 173.
 XX
 KW Human; gene therapy; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW proliferative disorder; inflammation; ss.
 OS Homo sapiens.
 XX
 PN WO200154474-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001349.
 XX

PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234977P.
 PR 25-SEP-2000; 2000US-0234988P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476161/51.

P-PSDB; ABB10285.

Isolated nucleic acid molecule encoding an inflammation-associated polypeptide is used in preventing, treating or ameliorating a medical

PT condition.

XX PS Claim 1; SEQ ID NO 173; 859pp + Sequence Listing; English.

XX CC The present invention provides human cDNAs, proteins and related genomic
 CC DNAs. These can be used in the treatment of neural, immune system,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders and inflammation. The present sequence
 CC is a cDNA of the invention

XX SQ Sequence 680 BP; 190 A; 138 C; 168 G; 183 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	1.06e-66	Length:	680
Score:	1278.00	Matches:	224
Percent Similarity:	99.56%	Conservative:	1
Best Local Similarity:	99.12%	Mismatches:	1
Query Match:	12.51%	Indels:	0
DB:	4	Gaps:	0

US-09-977-053-6 (1-1842) x ABA06507 (1-680)

Qy	1284	CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis	1303
Db	3	TGTGTTGATGGTGGCTGGCTATCGTGCCACATGTGGAAGGATTTGTAGCCCTGCAT	62
Qy	1304	CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu	1323
Db	63	TGTGAACAGAAAGTCAATGATGCCAGTCAACCCATGCTTAATAATGACAGTCTGTGA	122
Qy	1324	AspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly	1343
Db	123	GACCAAGTGGGGGATTCATGTGCAATGCCACCTGGATTTTGGTGACCGATGTGA	182
Qy	1344	LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValAlaThrCysLysAspGly	1363
Db	183	AGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	242
Qy	1364	AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn	1383
Db	243	GCCATAGCTTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	302
Qy	1384	IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn	1403
Db	303	ATCAATGAATGTCAGTCTTAATCATGTAGAAATCAGCCACCTGTGTGATGAAATTAAT	362
Qy	1404	SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer	1423
Db	363	TCTACAGTTGTAATGTCAGCAGGATTTTCAGGCAAGGTGTGAACAGAGTCT	422
Qy	1424	ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly	1443
Db	423	ACAGGCTTTAACCCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCATGCTAGATGC	482
Qy	1444	MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspMet	1463
Db	483	ATGCTCCCATCTCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	542
Qy	1464	AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuLeu	1483
Db	543	AACTATGACACCAATCTCTCTATGATGATGATGATGATGATGATGATGATGATGAT	602
Qy	1484	ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro	1503
Db	603	ACTGATTAACCGCTGGGTTCTTTATGATGATGATGATGATGATGATGATGATGATGAT	662
Qy	1504	SerValAsnAspGlyArg	1509
Db	663	TCGGTGAATGATGGCAGA	680

RESULT 23

AAS31580

ID AAS31580 standard; cDNA; 680 BP.

XX	AAS31580;
AC	04-DEC-2001 (first entry)
XX	cDNA encoding novel human calcium-binding protein #4.
DT	Human; calcium-binding protein; calcium flux; neurological disease;
XX	immune dysfunction; digestive disorder; neoplastic disease;
DB	blood disorder; infectious disease; gene therapy; immunosuppressive;
XX	antiarthritic; cytostatic; vasotropic; antibacterial; nootropic;
KW	virucide; ss.
XX	Homo sapiens.
OS	WO200155304-A2.
XX	02-AUG-2001.
PD	17-JAN-2001; 2001WO-US001302.
XX	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-018464P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217496P.
PR	14-JUL-2000; 2000US-0218290P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-022547P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226688P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.

```
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232403P.
PR 14-SEP-2000; 2000US-0232406P.
PR 14-SEP-2000; 2000US-0232405P.
PR 14-SEP-2000; 2000US-0232423P.
PR 21-SEP-2000; 2000US-0232427P.
PR 21-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251868P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
08-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465568/50.
P-PSDB; AAU19895.

Isolated nucleic acid molecule encoding a calcium-binding protein is used
in preventing, treating or ameliorating a medical condition.

Claim 4; SEQ ID NO 14; 542pp; English.

The present invention relates to the isolation of novel human calcium-
binding proteins (AAU19892-AAU19969), and cDNA and genomic sequences
encoding for these proteins. The sequences of the invention are useful in
the diagnosis, prevention and/or prognosis of diseases associated with
aberrant calcium flux. Such disorders include neurological diseases (e.g.
amyotrophic lateral sclerosis, ALS), immune dysfunction (e.g. severe
combined immunodeficiency, SCID), digestive disorders (e.g. irritable
bowel syndrome, IBS), neoplastic disease (e.g. cancer), blood disorders
(e.g. haemophilia), and/or infectious disease (e.g. acquired
immunodeficiency syndrome, AIDS). The novel calcium-binding proteins are
also useful as screening tools to identify antagonists and/or agonists
that may enhance or inhibit activities mediated by calcium-binding
proteins. The polynucleotides of the invention are also useful in gene
therapy. AAS31577-AAS31654 represent cDNA sequences encoding for the
novel human calcium-binding proteins. Note: The sequence data for this
patent did not form part of the printed specification, but was obtained
in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 680 BP; 190 A; 138 C; 168 G; 183 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1.06e-66 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 12.51% Indels: 0
DB: 4 Gaps: 0

US-09-977-053-6 (1-1842) x AAS31580 (1-680)

Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
Db 3 TGTGTTGATGTTGGCTGGCTATGTCACATGTTGAAGAGTTTGTAGCCCTGCAT 62

Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu 1323
Db 63 TGTGAACAGAGTCAATGAATGCCAGTCAAAACCCATGCTTAATAATGACAGTCTGTGAA 122

Qy 1324 AspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
Db 123 GACCAGGTGGGGGATTTCATGTGCAAAATGCCACCTGGATTTTGGTACCAGATGGA 182

Qy 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValAlaThrCysLysAspGly 1363
Db 183 AAGAACGTGATGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGCTACCTGTAAGACGGT 242

Qy 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
```

Db 243 GCCAATAGCTTCAGATGCTGTGTCAGCTGCTCAGGATCACACTGTGAATTGAC 302
Qy 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
Db 303 ATCAATGAATGTCACTCTAATCCATGATAGAAATCAGGCACCTGTGTGATGAATTAAT 362
Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlnSer 1423
Db 363 TCATACAGTTGTAAATGTTCAGCCAGATTTTCAGGCMAAAGGTGTGAACAGAACAGTCT 422
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyValIleTyrGlyTyrValMetLeuAspGly 1443
Db 423 ACAGCCTTTAATCTGATTTTGAAGTTCTGCACTATGATATGTCATGATAGTGC 482
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspMet 1463
Db 483 ATGCTCCCATCTCTCCATGCTCTAACCCTGATCTTCTGATGAATCTCTGACGACATG 542
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
Db 543 AACTATGGAACCAATCTCTATGATGATGATGATGATGATGATGATGATGATGATG 602
Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
Db 603 ACTGATTATAACGGCTGGGTTCTTTATGTAATGGCAGGAGAAAGATTAACAACCTGTC 662
Qy 1504 SerValAsnAspGlyArg 1509
Db 663 TGGTGAATGATGGCAGA 680
RESULT 24
ABT07797
ID ABT07797 standard; DNA; 680 BP.
XX AC ABT07797;
XX DT 14-NOV-2002 (first entry)
XX DE Novel human nucleic acid SEQ ID No 37.
XX KW Immunostimulant; antirheumatic; antiarthritic; neuroprotective;
KW KW antiallergic; antidiabetic; antiaesthetic; antiinflammatory; nootropic;
KW KW immunosuppressive; anticoagulant; thrombolytic; antiatherosclerotic;
KW KW cytosatic; nephrotropic; antiparkinsonian; gynecological; virucide;
KW KW antibacterial; antiarrhythmic; fungicide; HCFAT05; HMAAE95; HTNBM01;
KW KW immunodeficiency; autoimmune disorder; allergic reaction; cardiovascular;
KW KW inflammatory condition; graft-versus-host disease; reproductive system;
KW KW blood-related disorder; hyperproliferative; endocrine; neurological;
KW KW respiratory; renal; infectious disease; gastrointestinal; gene therapy;
KW KW neuronal growth; neuronal disorder; neuro-degenerative condition;
KW KW keratinocyte growth; human; ds.
OS Homo sapiens.
XX XX
XX XX
XX XX US2002086330-A1.
XX XX
XX XX PD 04-JUL-2002.
XX XX
XX XX PP 17-JAN-2001; 2001US-00764893.
XX XX
XX XX PR 31-JAN-2000; 2000US-0179065P.
XX XX PR 04-FEB-2000; 2000US-0180628P.
XX XX PR 28-JUN-2000; 2000US-0214886P.
XX XX PR 07-JUL-2000; 2000US-0216647P.
XX XX PR 07-JUL-2000; 2000US-0216880P.
XX XX PR 11-JUL-2000; 2000US-0217487P.
XX XX PR 11-JUL-2000; 2000US-0217496P.
XX XX PR 14-JUL-2000; 2000US-0218290P.
XX XX PR 26-JUL-2000; 2000US-0220963P.
XX XX PR 26-JUL-2000; 2000US-0220964P.
XX XX PR 14-AUG-2000; 2000US-0224518P.
XX XX PR 14-AUG-2000; 2000US-0224519P.
XX XX PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-023935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX XX (ROSE/) ROSEN C A.
XX XX (RUBE/) RUBEN S M.
XX XX (BARA/) BARASH S C.
XX XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2002-665432/71.
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
XX treatment of immune, hyperproliferative, renal, respiratory,
XX cardiovascular, reproductive, endocrine, gastrointestinal and
XX neurological disorders.
XX XX
XX PS Disclosure; Page 226; 335pp; English.
XX XX
XX CC The invention relates to an isolated polypeptide comprising a sequence at
XX CC least 90% identical to a full length protein sequence selected from 55
XX CC sequences given in the specification such as a sequence of 163, 74 or 140
XX CC amino acids fully defined in the specification, or the encoding sequence
XX CC contained in 49 cDNA clones given in specification e.g. HCFAT05, HMAAE95
XX CC or HTNBM01. The protein and its encoding nucleic acid are useful for
XX CC diagnosing a pathological condition or susceptibility to a pathological
XX CC condition in a subject and for preventing, treating or ameliorating a
XX CC medical condition. The protein, its encoding nucleic acid and an isolated
XX CC antibody that can bind to the protein are useful in treating, preventing,
XX CC diagnosing and/or prognosing immunodeficiencies, autoimmune disorders,
XX CC allergic reactions and conditions, inflammatory conditions, graft-versus-
XX CC host disease, blood-related disorders, hyperproliferative disorders,
XX CC renal disorders, cardiovascular disorders, respiratory disorders,
XX CC neurological disorders, endocrine disorders, reproductive system
XX CC disorders, infectious diseases, and gastrointestinal disorders. The
XX CC protein of the invention is useful to stimulate neuronal growth and to
XX CC treat, prevent, and/or diagnose neuronal damage which occurs in certain
XX CC neuronal disorders or neuro-degenerative conditions, for stimulating
XX CC keratinocyte growth, to prevent hair loss, to modulate mammalian

PA (RUBE/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2002-681727/73.
 DR P-PSDB; ABP66872.
 XX
 PT Novel polypeptide useful for diagnosis, prognosis, prevention, and
 PT treatment of immune, hyperproliferative, renal, respiratory,
 PT cardiovascular, reproductive, endocrine, gastrointestinal and
 PT neurological disorders.
 XX
 PS Claim 1; SEQ ID NO 173; 369pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABV83682-ABV84101) and proteins
 CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC the nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 680 BP; 190 A; 138 C; 168 G; 183 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 1,06e-66 Length: 680
 Score: 1278.00 Matches: 224
 Percent Similarity: 99.56% Conservative: 1
 Best Local Similarity: 99.12% Mismatches: 1
 Query Match: 12.51% Indels: 0
 DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x ABV83844 (1-680)

Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
 Db 3 TGTGTTGATGTTGGCTGCTATCGTTGCACATGTTGAAAGGATTTGTAGGCTGCAT 62

Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
 Db 63 TGTGAACAGAGTCAATGATGCCAGTCAACCCATGCTTAAATATGTCAGTCTGTGAA 122

Qy 1324 AspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
 Db 123 GACCAGGTTGGGGATTCATGTGCAATGCCACCTGGATTTTGGGTACCCGATGGA 182

Qy 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
 Db 183 AAGAACGTGCGATGATGTTCTCAGTCAGCCATGCCAAAATGGAGTACCTGTAAGACCGT 242

Qy 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
 Db 243 GCCAATAGCTTCAGATGCTGTGTGTCAGCTGGCTTCACAGGATCACACTGTGAATTGAAC 302

Qy 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
 Db 303 ATCAATGATGTCACTTAAATCCATGTAGAAATCAGGCCACCTGTGTGGATGAATTAAT 362

Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
 Db 363 TCATACAGTTGTAAATGTTCAGCCAGGATTTTTCAGGCAAGGTGTGAAACAGACAGTCT 422

Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
 Db 423 ACAGGCTTTAACCTGGATTTTGAAGTTTCTGCACTATGGATATGTCATGCTAGATGGC 482

Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSerSerAspMet 1463
 Db 483 ATGCTCCCATCTCTCCATGCTCTAACCTGTACCTTCTGGATGAAATCTCTGACGACATG 542

Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
 Db 543 AACTATGGAACACCAATCTCTATGAGTTGATAACGGCAGGACAAATACCTTGTCTCTG 602

Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
 Db 603 ACTGATTATACGGCTGGTCTTTTATGTGATGGCAGGAGAAAGATACAAACTGTCTCC 662

Qy 1504 SerValAsnAspGlyArg 1509
 Db 663 TCGGTGAATGATGGCAGA 680

Search completed: May 9, 2004, 16:20:25
 Job time : 1579.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 10, 2004, 15:48:01 ; Search time 265.768 Seconds

(without alignments)
3846.286 Million cell updates/sec

Title: US-09-977-053-6

Perfect score: 10219

Sequence: 1 MWPLAFCCGLALVSGWAT.....GVTKICLBSGEWHILPVC 1842

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_P/US09977053/runat_06052004_075943_18213/app_query.fasta_1.5710
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -NORM=ext -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=25
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977053 &C&N 1 420 @runat_06052004_075943_18213 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10208	99.9	10878	4	US-09-911-842A-1
2	8652.5	84.7	11230	4	US-09-911-842A-3
3	914.5	8.9	9723	1	US-08-083-590A-21
4	914.5	8.9	9723	3	US-08-532-384-21
5	838	8.2	763	4	US-09-484-970B-110
6	829	8.1	8091	4	US-09-230-652-1
7	829	8.1	8257	4	US-09-484-970B-65
8	764.5	7.5	6677	4	US-08-939-366-27
9	764.5	7.5	6677	4	US-09-467-597-6
10	749	7.3	4208	3	US-09-214-278-6
11	749	7.3	4208	4	US-09-855-722-6
12	745	7.3	5458	4	US-09-199-865-2

13	745	7.3	6464	2	US-08-400-159-5	Sequence 5, Appli
14	745	7.3	6464	3	US-08-611-729A-5	Sequence 5, Appli
15	744	7.3	4208	4	US-09-068-740A-10	Sequence 10, Appli
16	744	7.3	5590	3	US-08-882-046-1	Sequence 1, Appli
17	690	6.8	3582	2	US-08-400-159-9	Sequence 9, Appli
18	690	6.8	3582	3	US-08-611-729A-9	Sequence 9, Appli
19	674	6.6	3955	3	US-09-214-278-4	Sequence 4, Appli
20	674	6.6	3955	4	US-09-855-722-4	Sequence 4, Appli
21	671	6.6	4464	2	US-08-400-159-7	Sequence 7, Appli
22	671	6.6	4464	3	US-08-611-729A-7	Sequence 7, Appli
23	642.5	6.3	4315	3	US-08-882-046-3	Sequence 3, Appli
24	633	6.2	2508	3	US-08-981-392-1	Sequence 1, Appli
25	633	6.2	2508	3	US-08-981-392-3	Sequence 3, Appli
26	631.5	6.2	2663	4	US-09-068-740A-8	Sequence 8, Appli
27	631.5	6.2	2663	4	US-09-423-753-8	Sequence 8, Appli
28	631	6.2	2692	3	US-08-981-392-11	Sequence 11, Appli
29	620.5	6.1	2892	1	US-08-264-534-5	Sequence 5, Appli
30	620.5	6.1	2892	1	US-08-083-590A-1	Sequence 1, Appli
31	620.5	6.1	2892	1	US-08-465-500-5	Sequence 5, Appli
32	620.5	6.1	2892	2	US-08-346-126-5	Sequence 5, Appli
33	620.5	6.1	2892	2	US-08-346-128-5	Sequence 5, Appli
34	620.5	6.1	2892	3	US-08-532-384-1	Sequence 1, Appli
35	620.5	6.1	2892	3	US-08-893-828-5	Sequence 5, Appli
36	606	5.9	5561	2	US-08-400-159-1	Sequence 1, Appli
37	606	5.9	5561	3	US-08-611-729A-1	Sequence 1, Appli
38	602	5.9	8378	5	PCT-US91-09055-1	Sequence 1, Appli
39	594	5.8	5176	4	US-09-182-024A-1	Sequence 1, Appli
40	592	5.8	2857	3	US-08-981-392-4	Sequence 4, Appli
41	587	5.7	3231	4	US-09-866-028-14	Sequence 14, Appli
42	585	5.7	2055	3	US-08-872-855-3	Sequence 3, Appli
43	585	5.7	2800	3	US-08-872-855-1	Sequence 1, Appli
44	585	5.7	3339	4	US-09-423-753-4	Sequence 4, Appli
45	583.5	5.7	5583	4	US-09-312-283C-372	Sequence 372, App
46	572	5.6	4758	3	US-09-191-647-1	Sequence 1, Appli
47	572	5.6	4758	3	US-09-540-245A-1	Sequence 1, Appli
48	572	5.6	4758	3	US-09-540-153-1	Sequence 1, Appli
49	549.5	5.4	2481	3	US-09-188-930-64	Sequence 64, Appli
50	549.5	5.4	2481	4	US-09-312-283C-64	Sequence 64, Appli
51	549.5	5.4	6951	4	US-09-023-655-1265	Sequence 1265, Ap
52	548.5	5.4	6951	6	5256642-1	Patent No. 5256642
53	548.5	5.4	6951	6	5472939-1	Patent No. 5472939
54	546.5	5.3	7313	6	5472939-3	Patent No. 5472939
55	525.5	5.1	5420	6	5256642-3	Patent No. 5256642
56	525.5	5.1	5420	6	5472939-3	Patent No. 5472939
57	502	4.9	5499	3	US-08-479-722B-1	Sequence 1, Appli
58	502	4.9	5502	5	PCT-US95-02251-17	Sequence 17, Appli
59	499.5	4.9	1981	3	US-08-981-392-26	Sequence 26, Appli
60	471.5	4.6	11272	4	US-09-341-461-1	Sequence 1, Appli
61	467	4.6	1434	4	US-09-833-381-926	Sequence 926, App
62	454	4.4	2899	3	US-08-981-392-24	Sequence 24, Appli
63	452.5	4.4	5089	6	517197-31	Patent No. 517197
64	429.5	4.2	1556	1	US-08-597-545-3	Sequence 3, Appli
65	429.5	4.2	1556	1	US-08-457-135-3	Sequence 3, Appli
66	427.5	4.2	1158	4	US-09-142-027A-11	Sequence 11, Appli
67	427.5	4.2	1158	4	US-09-142-027A-13	Sequence 13, Appli
68	427.5	4.2	1573	1	US-08-597-545-4	Sequence 4, Appli
69	427.5	4.2	1573	1	US-08-457-135-4	Sequence 4, Appli
70	406.5	4.0	1161	4	US-09-833-381-2046	Sequence 2046, Ap
71	395.5	3.9	3142	1	US-08-110-158-3	Sequence 3, Appli
72	395.5	3.9	3142	4	US-09-023-655-1090	Sequence 1090, Ap
73	395.5	3.9	3144	5	PCT-US91-05059-1	Sequence 1, Appli
74	390	3.8	1785	4	US-09-148-545-112	Sequence 112, App
75	389.5	3.8	3759	3	US-08-479-722B-3	Sequence 3, Appli
76	382.5	3.7	6942	2	US-08-460-309-3	Sequence 3, Appli
77	382.5	3.7	6942	2	US-08-125-077-3	Sequence 3, Appli
78	382.5	3.7	9534	4	US-09-562-702A-5	Sequence 5, Appli
79	382.5	3.7	9534	4	US-09-561-709B-8	Sequence 8, Appli
80	382.5	3.7	9535	4	US-09-562-702A-1	Sequence 1, Appli
81	378.5	3.7	9419	4	US-09-562-702A-7	Sequence 7, Appli
82	378.5	3.7	9420	4	US-09-562-702A-3	Sequence 3, Appli
83	377.5	3.7	3934	4	US-09-023-655-1066	Sequence 1066, Ap
84	372.5	3.6	1448	4	US-09-130-491-9	Sequence 9, Appli
85	368	3.6	5197	4	US-09-833-381-1910	Sequence 1910, Ap

86	362.5	3.5	1117	4	US-09-148-545-54	Sequence 54, Appl
87	362.5	3.5	9391	4	US-09-562-702A-11	Sequence 11, Appl
88	362.5	3.5	9511	4	US-09-563-702A-9	Sequence 9, Appl
89	360	3.5	3753	5	PT-US95-02251-2	Sequence 2, Appl
90	360	3.5	4314	1	US-08-199-780-2	Sequence 2, Appl
91	360	3.5	4314	2	US-08-316-650-2	Sequence 2, Appl
92	359	3.5	3449	4	US-09-016-434-1453	Sequence 1453, Ap
93	359	3.5	3449	4	US-09-023-655-1478	Sequence 1478, Ap
94	358	3.5	1592	3	US-08-252-493C-1	Sequence 1, Appl
95	358	3.5	1592	3	US-09-276-197-1	Sequence 1, Appl
96	358	3.5	1833	1	US-08-365-470-2	Sequence 2, Appl
97	358	3.5	3834	1	US-09-209-668-18	Sequence 18, Appl
98	358	3.5	3854	1	US-08-365-470-1	Sequence 1, Appl
99	358	3.5	3858	4	US-08-344-155C-98	Sequence 98, Appl
100	358	3.5	3858	4	US-09-009-490A-58	Sequence 88, Appl

ALIGNMENTS

```

RESULT 1
US-09-911-842A-1
; Sequence 1, Application US/09911842A
; Patent No. 6656707
; GENERAL INFORMATION:
; APPLICANT: Amgen Inc.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCES: 01017737592
; CURRENT APPLICATION NUMBER: US/09/911,842A
; CURRENT FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 10878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-911-842A-1

```

Alignment Scores:		
Pred. No.:	0	Length: 10878
Score:	10208.00	Matches: 1838
Percent Similarity:	99.95%	Conservative: 3
Best Local Similarity:	99.78%	Mismatches: 1
Query Match:	99.89%	Indels: 0
DB:	4	Gaps: 0

US-09-977-053-6 (1-1842) x US-09-911-842A-1 (1-10878)

Qy	1	MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr	20
Db	11	ATGTGGGCTCGCCTGGCCTTTTGTGCTGGGCTCTGGCGCTCGGCTTGGGCTGGGCGACC	70
Qy	21	PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro	40
Db	71	TTTCAGCAGATGTCCCGTCGCGAAATTCAGCTTCGCTCTTCCCGAGACCCGCGCC	130
Qy	41	GlyAlaProGlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArg	60
Db	131	GGGGCCCCCGGGGATCCCCCGCGCGCGCTCTGTGGCGACGAAGCGCGGGAGCAGA	190
Qy	61	ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu	80
Db	191	GTGAGAGCGGCTGGGCGAGGCGTTCCGGCGACGCGTCCGGCTCTCGGGGAGCTCAGCGAG	250
Qy	81	ArgLeuGluLeuValPheLeuValAspAspSerSerSerValGlyGluValAsnPheArg	100
Db	251	GGCTTGGAGCTTGTCTCTCGTGGATGATCGTCCAGCGTGGCGGAGTCAACTTCGCG	310
Qy	101	SerGluLeuMetPheValArgLysLeuSerAspPheProValValProThrAlaThr	120
Db	311	AGCCAGCTCATGTTCGTCCGAAGCTGTGTCCGACTTCCCGTGGTGCCACGCGCAGC	370

1451 CAAGGAAACAGCAGTGGGATGGCCAGAACCCGGTGTGTGGAGGCCACTGTTCCACC 1510
Qy PheGlnMetProIysAspValIlelleSerProHisAsnCysGlyLysGlnProAlaLys 520
1511 TTTTCAGATGCCCAAGATATCATATCATATCCCCCAACACTGTGGCAAGCAGCCAGCCAAA 1570
Qy PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
1571 TTTGGAGCATCTGTATGTAAGTGTGCCCAAGGGTTCATTTTATCTGGAGTCACAGAA 1630
Qy MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
1631 ATGCTGAGATGTACCACTTCTGGAAATGGAATGTGGAGTTCCAGGCAGCTGTGTGTAAA 1690
Qy AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGln 580
1691 GAGGTGGAGGCTCTCAAAATCACTGTCTTAAGACATAGAGCTTAAGACTCTGGAAACAG 1750
Qy GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
1751 CAAGATTCTGCCATGTACCTGGAGATTCGAAGCTTCAAGACACTCTGGTGAAGAG 1810
Qy ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
1811 GTGTCAGTCCAGTTCATCCAGCTTTCCAGCCCACTTACCTTTTCCAGCTGGAGATGT 1870
Qy AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
1871 GCTATCGTATACAGCGCAACTGACCTATCCGGCAACCGCCAGCTGCACTTTTCCATATC 1930
Qy LysValIleAspAlaGluProValIleAspTrpCysArgSerProProValGln 660
1931 AAGGTATTGATGCAGAACCACTGTCNTAGACTGTGTGCAGATCTCCACCTCCGCTCCAG 1990
Qy ValSerGluLysValHisAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
1991 GTCTCGGAGAGGTACATGCGCGAGCTGGATGAGCCTCAGTCTTCAGACAACTCAGGG 2050
Qy AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
2051 GCTGAATTGGTCTATTACCAAGAGTCTATACACAGGAGACCTTTTCCCTCAAGGGAGACT 2110
Qy IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
2111 ATAGTACAGTATACGCCACTGACCTCCTCAGGCATTAACAGGACATGTGATATCCATAT 2170
Qy ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
2171 GTCATAAAAGGTCTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATATGC 2230
Qy ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
2231 ACTCCAGATAATCTGGAGTCACATGTACATTAATCTGTGGAGGCTTAATGATTTCCACA 2290
Qy GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
2291 GAAGGTCTACTGACAGATTAATTTGTGCTTATGAGATGGGTCTGGAAACCCACATAT 2350
Qy ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
2351 ACCACTGAATGGCCAGACTGTGCGCAAAAACGTTTGTCTAACCAACCGGTTTCAAGTCTCTT 2410
Qy GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
2411 GAGATGTTCTACAAAGCAGCTCGTGTGTGATGACACAGATCTGATGAAGAAGTTTCTGAA 2470
Qy AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
2471 GCATTTGAGACGACCTCTGGAAAAAATGGTCCCATCATTTTGTAGTGAAGAGGACATT 2530
Qy AspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860

2531 GACTGCAGACTGGAGAGAACCTGACCAAAAAATATTTCCTAGATAATAAATTATGACTAT 2590
Qy GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAlaAsnArgLeuAspTyrSer 880
2591 GAAATGGCTTTTGCATTTGGACCGAGTGGCTGGGTGCAGCTAATAGCTGGATCTCT 2650
Qy TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
2651 TAGCATGACTTCTCTGCAGACTGTGCAAGAAACAGCCACAGCATGGCCAAATGCCAGTCC 2710
Qy SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
2711 TCACGATTAATAAAGAGTGGCCCATTTATCTGACTATATAAATTAAGTTAATTTTAAACATC 2770
Qy ThrAlaSerValProLeuProAspGluTrpAsnAspThrLeuGluTrpGluAsnGlnGln 940
2771 ACAGCTATGTGTGCATTTACCCGATGAAGAATGATACCTTTGATGGGAANAATCAGCA 2830
Qy ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
2831 CGACTCCTTCAGACATTTGGAACTATCACAAATAAATGAAAGGACTCTCAACAAGAC 2890
Qy ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
2891 CCCATGTTATTCCTTTTCAGCTTGCATCAGAAATATCTATAGCCGACGCAATTTCAATGAA 2950
Qy ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
2951 ACAAAGAAGCTTTCCCTTCTGCAGACCAAGCTCAGTGTGAGAGGGCGTATGTGTGTC 3010
Qy AspCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
3011 AATTGCCCTTTGGAACTATTATATCTGGAACATTTCACTGTGAAAGCTGCGGATC 3070
Qy GlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
3071 GGATCTCTATCAAGTGAAGAAGGCGCACTTGAGTGCAGAGCTTTGCCCTCTGGGATGTAC 3130
Qy ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
3131 ACGAATATATCATTTCAAGAAACATCTGATTTGTAAGCTCAGTGTAAACAGGACAC 3190
Qy TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
3191 TACTCATACAGTGGACTTGAGACTTGGAATCGTGTCCACTGGGCACCTTATCAGCCAAA 3250
Qy PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
3251 TTTGGTTCCCGAGCTGCCCTCTCGTGTCCAGAAACACCTCAACTGTGAAAGAGGAGCC 3310
Qy ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
3311 GTGACATTTCTGCATGTGAGTCTCTTGTCCAGAGGAAATTTCTCGGTTCTGGGTTA 3370
Qy MetProCysHisProCysProArgAspTyrTyrGlnProAlaGlyLysAlaPheCys 1140
3371 ATGCCCTGTCAACCATGTCTCTGTGACTATTACCACTTAATGAGGAGGAGGCTTCTGC 3430
Qy LeuAlaCysProPheTyrGlyThrThrPheAlaGlySerArgSerIleThrGluCys 1160
3431 CTGCGCTGTCCCTTTTATGGAATCTACCCCATTCGCTGTTCAGATCCATCAGAAATGT 3490
Qy SerSerPheSerThrPheSerAlaAlaGluSerValValProProAlaSerLeu 1180
3491 TCAAGTTTATGTTCAACTTTCTCAGCGCAGAGGAAAGTGTGGTGGCCCTCTCTCT 3550
Qy GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
3551 GGACATATTAAGAAGAGCATGAATCAGCAGTCAAGTGTTCATGATGCTCTTTTAAAC 3610
Qy ProCysHisAsnSerGlyThrCysGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
3611 CTTTGCACAAATAGTGGAACTTGCAGCAACTTGGGCGTGGTATATGTTGTCTCTGTCCA 3670

Qy	1221	LeuGlyTyrThrGlyLeuLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys	1240
Db	3671	CTTGGATATACAGGCTTAAAGTGTGAACAGACATCATGATGAGTGACGCCACCTGCCTTGC	3730
Qy	1241	LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly	1260
Db	3731	CTCAACAATGGAGTTTGTAAAGACCTAGTTGGGGAAATTCATTTGTGAGTGCCCATCAGGT	3790
Qy	1261	TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn	1280
Db	3791	TACACAGGTGAGCGGTGTGAAGAANAATATAAATGAGTGTAGCTCCAGTCTCTTGTTTAAAT	3850
Qy	1281	LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal	1300
Db	3851	AAAGCAATCTGTGTTCATGGTGTGGCTGGCTATCGTTGTCACATGTGTGAAGGATTTGTA	3910
Qy	1301	GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAla	1320
Db	3911	GGCTCGATATGTGAACAGAGAGTCAATGAATGCCAGTCAAAACCCATGCTTTAAATAAAGCA	3970
Qy	1321	ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThr	1340
Db	3971	GTCTGTGAGACACAGTTGGGGATTTCTTGTGCAAAATGCCACCTGATTTTGGGTACC	4030
Qy	1341	ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys	1360
Db	4031	CGATGTGGAAGAAACGTGATGAGTGTCTCAGTCAGCCATGCCAAAAATGGAGCTACCTGT	4090
Qy	1361	LysAspGlyAlaAsnSerPheArgCysLeuCyValaAlaGlyPheThrGlySerHisCys	1380
Db	4091	AAAGCGGTGCAATAGCTTCAGATGGCTGTGTGCAGTGCCTTCTACAGAGTACACTGT	4150
Qy	1381	GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp	1400
Db	4151	GAATTGAACATCAATGAATGTCTAGTCTAATCCATGTAGAAATCAGGCCACCTGTGTGGAT	4210
Qy	1401	GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr	1420
Db	4211	GAATTAATAATTATACAGTTGTAAATGTTCAGCCAGGATTTTTCAGCAAAAAGGTGTGAACA	4270
Qy	1421	GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet	1440
Db	4271	GAACAGTCTACAGGCTTTAACCTGGATTTTGAAGTTTCTGGCATCATATGATATGTCATG	4330
Qy	1441	LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer	1460
Db	4331	CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACCTGTACTTCTGGATGAATCCCTCT	4390
Qy	1461	AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr	1480
Db	4391	GACGACATGAACATATGGAACACCAATCTCTCATGAGTTGATATACGCGACGCAATACC	4450
Qy	1481	LeuLeuLeuThrAspTyrAsnGlyTyrPheValLeuTyrValAsnGlyArgGluLysIleThr	1500
Db	4451	TTGCTCTCTGATGATTAACCGCTGGGTCTTTTATGTGAAATGGCAGGGAAGAATAACA	4510
Qy	1501	AsnCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAla	1520
Db	4511	AACTGTCCCTCGGTAATGATGGCAGATGGCATCATATTCATCTTGGATGAATCCCTCT	4570
Qy	1521	AsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyAlaGlyLeuSer	1540
Db	4571	AATGGCATCTGGAAGCTATATATCGATGGGAATTAATCTACGCGGTGGTGTGGCTCTCT	4630
Qy	1541	ValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLys	1560
Db	4631	GTGCTTGGTCCCATACTCTGGTGGTGGTCCGTAGTCTCTGGGCGAGAGCAACAAAAA	4690
Qy	1561	GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp	1580
Db	4691	GGAGAGGATTCAGCCAGCTGAGTCTTTTGTGGGCTCCATAAGCAGCTCAACCTCTGG	4750

RESULT 2

```

US-09-911-842A-3
; Sequence 3, Application US/09911842A
; Patent No. 6656707
; GENERAL INFORMATION:
; APPLICANT: Amgen Inc.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911,842A
; CURRENT FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 11230
; TYPE: DNA

```

ORGANISM: Mus musculus

US-09-911-842A-3

Alignment Scores:

Pred. No.: 0 Length: 11230
Score: 8652.50 Matches: 1532
Percent Similarity: 90.56% Conservative: 137
Best Local Similarity: 83.13% Mismatches: 171
Query Match: 84.67% Indels: 3
DB: 4 Gaps: 3

US-09-977-053-6 (1-1842) x US-09-911-842A-3 (1-11230)

QY 1 MetTyrProArgLeuAlaPheCysGlyTyrGluLeuAlaValSerGlyTyrAlaThr 20
DB 176 ATGTGTGTCGCGCTGTCCTTTTGTGTCGGCTCTGGCACTGTGTGCGGCTGGACCAAC 235
QY 21 PheGlnGlnMetSerProSerArgGlnPheSerPheArgLeuPheProGluThrAlaPro 40
DB 236 TTCACGCGTGGCCCTTGGCTCACTTCAGCTTCGGCTGTCCCGAGGCTCTCCG 295
QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
DB 296 GGGCTCTGGCAGACTGGCGGTACCTCCCGCTCCAGTGAGGAGGAGGAGGAGGAGC 355
QY 60 ArgValGluArgLeuGlnAlaPheArgArgValArgValArgLeuLeuLeuSer 79
DB 356 AAGTGGAGCGCTGGCGCGCGCTTCGCGAGCGCGTGGCGGAGCTCAGC 415
QY 80 GluArgLeuGluLeuValPheLeuValAspSerSerValGlyGluValAlaPhe 99
DB 416 GCGAGCTGGAGCTGCTTCTCTGGTGGAGAGTGGTCCAGCGTGGCCAAACCACTTC 475
QY 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAla 119
DB 476 CTCACGAGCTCAAGTTCGTGGCAAGCTGCTGTCCGACTTCCCGTGTGTCCACGGCC 535
QY 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
DB 536 ACGGTGTGGCATCGTCACTTCTCATCCAAAGCAACAGTGTGTGGCGGTGGATTC 595
QY 140 IleSerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAla 159
DB 596 ATCTCCACGAGCGCGCGCACCAACACAGTGGCGCTGTCTCAGCGCGGAGATCCCGGC 655
QY 160 IleSerThrArgGlyGlyThrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
DB 656 ATCACTACCGCGTGTGGCACTATACCAAGGGCGCTTCCAGCAAGCGCGCAATC 715
QY 180 LeuLeuHisAlaArgGluLeuSerThrLysValValPheLeuIleThrAspGlyTyrSer 199
DB 716 CTTCGTCACTCTAGAGAAACTCCCAAGTCAATATTTCTCATCCGCGGCTATTC 775
QY 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
DB 776 AATGGCGAGACCCAGACCTATTGCGCATCGCTTCGGGATTTCCGAGTGGAGATCTTC 835
QY 220 ThrPheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
DB 836 ACGTTCGGGATTTGGCAGGGAATATCCGGGAATCGAATGACATGGCTTCACCCCGAAG 895
QY 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheLeuAlaLeuAlaArgAla 259
DB 896 GAAGAACATTTGTACCTGCTCCACAGTTTGAAGAAATTTGAGGCTTTAGCTCGAGGGCG 955
QY 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
DB 956 TTGCATGAAGATCTACTCTCTGGAGTTTATCCAGAGATATGGCCCACTCTCTAT 1015
QY 280 LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis 299
DB 1016 CTCTGTGAGGCTGGGAAGACTGCTGTGACAGAAATGGCCAGCTGCCAAATGTGGGACACAC 1075

QY 300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGlu 319
DB 1076 ACGGTCAATTTGATGCACTCTGAGAGGCTATTACGGGAAGGCTCTGAGCATGAG 1135
QY 320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
DB 1136 TGCACAGCTTGGCCATCAGGCACATATAAGCCGAGCTTCTCCAGGAGGAATCAGCACC 1195
QY 340 CysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAsp 359
DB 1196 TGCATCCCATGTCCTGACGTAAAGCCACACTCCCACTGGAAGCACTTCCCTCTGAAGAC 1255
QY 360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
DB 1256 TGTGTGCGAGAGGATCCAGAGATCTGGCCAGACCTGTGAGTGTGTTCACCTGTCT 1315
QY 380 AlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
DB 1316 GCCTGAAGCTCTCTGAATAATGTTTATATACAAACACTTCGCAAAACACTTCTCAAT 1375
QY 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeu 419
DB 1376 GCGCCTGTGGGTCCGATGTGCGCCGCGCTTTGACCTTGTGGGAAGCAGCATTCATTG 1435
QY 420 CysLeuProAsnGlyLeuTyrSerGlySerGlySerTyrCysArgValArgThrCysPro 439
DB 1436 TGTCAACCCAAATGTTTGTGTGTGGAGCAGAAAGCTTCTGCAGAGTGAAGACGTGCC 1495
QY 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys 459
DB 1496 CACTCCGACAGCCCAACACAGCCACATCAGCTGCTCCACTGCGGAATGTCTCTAAC 1555
QY 460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr 479
DB 1556 ACCTGTGTTTGTGTACCTGAAGATGAGATTTAGAAAGCAGCAGCTAGGCTTACC 1615
QY 480 CysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSer 499
DB 1616 TGTCAAGGAATGCGCAGTGGATGGCCAGAGCCCGGTGTGTAGAACGCCATTTGTGCC 1675
QY 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
DB 1676 ACCTTCAGAAAGCCCAAGAGCGTCATCTTCTCACCCAGCTGCGGAGCAGCGCCGCC 1735
QY 520 LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
DB 1736 AGCCTGGGATGACCTGTGAGTAAAGCTGCGCAGGAGATACATTTTATCCGGGTGAGA 1795
QY 540 GluMetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAlaValCys 559
DB 1796 GAA---GTGAGATGTGCCACATCTGGGAAGTGGAGTGCACAAAGTTCAGACAGCTGTGTC 1852
QY 560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu 579
DB 1853 AAGATGTGGAGGCTCCACAAATCAGCTGTCCAAATGACATTTAGGCAAGACTGGGAG 1912
QY 580 GlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyGlu 599
DB 1913 CAGCAGACTCTGCTAATGTCACCTGGCAAGTCCCAACAGCTAAAGACAACTCTGGTGA 1972
QY 600 LysValSerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAsp 619
DB 1973 AAGGTGTAGTCCAGCTCCACCCAGCCTTTACCCCACTTACCTCTTCCCAATTTGGAGAC 2032
QY 620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
DB 2033 GTGGCCATCACTTACAGCGCAACCCACTCATCCGTTAAACCAAGCAGCTGCTCTTCTAC 2092
QY 640 IleLysValIleAspAlaGluProProValIleAspTyrCysArgSerProProVal 659
DB 2093 ATTAAGGTCTATTGATGTGGAACCGCTGTCTATAGATTGGTGCCGATCTTCCACCTCCAATC 2152
QY 660 GlnValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSer 679


```
Qy 933 rLeuGluTrpGluAsnGlnGlnArgLeuLeu-----GlnThrLeuGluThrI 949
Db 1685 -----CAGGTTTCACCTGGTGTGTGTGTCAGGAGAACATTGACAACTGT 1728
Qy 949 eThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAla 969
Db 1729 GACCCC-----GATCCTTGGCACCACCATGTCAGTGTCCAGGA 1763
Qy 969 rGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerProPheCysAr 989
Db 1764 TGTATT-----GATTCCTACACCTGC-----ATCTGCAA 1793
Qy 989 gProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTrAs 1009
Db 1794 TCCGGGTACATG-----GGGGCCATCTGCAGTGAC--CAGATTCAATGAATGTACAG 1844
Qy 1009 nLeuGluHisPheThrCysGluSerCys-----ArgIleGlySerTyrGlnAspGluI 1027
Db 1845 CAGCCCTTGCCTGAACGATGGTGCCTGCATTGACCTGGTCAATGGCTAC----- 1893
Qy 1027 uGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerAr 1047
Db 1894 -----CAGTGCACACTGCCAGCCAGCCAGCTCAGGGGTTAATTGTGAATT----- 1938
Qy 1047 gAsnIleSerAspCysLysAlaGln-----CysLysGlnGlyThrTy 1061
Db 1939 -AATTTTGATGACTGTGCAAGTAACCTTGTATCCATGGAATCTGTATGATGTCATTAA 1997
Qy 1061 rSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPh 1081
Db 1998 TCGCTACAGT-----TGTGTCTGTCTACACAGGATTCACA----- 2031
Qy 1081 eGlySerArgSerCysLeuSerCysProGluAsnThrSerThr---ValLysArgGlyAl 1100
Db 2032 -GGGCAGAGATGTAACATTGACATTGATGAGTGTGCTCCATCCCTGTCGACAGGTC 2090
Qy 1100 aValAsnIleSerAlaCysGlyValPro-----CysProGluGlyLysPh 1115
Db 2091 AACATGTATCAAC-----GGTGTGAATGGTTCCGGCTGTATATGCCCGAGGGA----- 2139
Qy 1115 eSerArgSerGlyLeuMetProCysHisPro-----CysProArgAspTyrTyrGlnProAs 1134
Db 2140 -----CCCCATCACCCAGCTGC-----TACTCACAGGTGAA 2171
Qy 1134 nAlaGlyLysAlaPheCysLeuAlaCysPro---PheTyrGlyThrThrProPheAlaG 1153
Db 2172 CGAA-----TGCCTGAGCAATCCCTGCATCCATGGA----- 2202
Qy 1153 ySerArgSerIleThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSe 1173
Db 2203 -----AACTGTACTGAGGTCTCAGTGGATATATAGTGTCTC----- 2238
Qy 1173 rValValProProAlaSerLeuGlyHisIleLysLysArgHisGluLysSerSerGlnVa 1193
Db 2239 -----TGTGATGAGGCTGGTGGTCATCAACTGTGAGGAACTTGTGACAAATCTGGTGA 2280
Qy 1193 lPheHisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyAr 1213
Db 2281 -----AATGAATGCCCTTTCGAATCCATCCAGAAATGGAGGAACTTGTGACAAATCTGGTGA 2336
Qy 1213 gGlyTyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAs 1233
Db 2337 TGAATACAGGTGACTTACTGCAAGAGGCTTTAAAGGCTATAAATGCCAGGTGAATATGA 2396
Qy 1233 pGluCysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPh 1253
Db 2397 TGAATGCTCAAAATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 2456
Qy 1253 eIleCysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGly 1273
Db 2457 CACTTGCCACTGTGTGTGTCATACACAGCAAGAAATTGTGCACAGTATTGGTCCCTG 2516
```

	: : :	3376 GGT--GTCGCAATCACTGGCCAACAGCATTA CTGTCAGTGC CCC---CTGGGCTAT 3429
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1637 SerValProHisLeu-----ArgThrAlaSerGluAspLeuLysProGly 1651
Q _y		
	: : : : : : : :	3430 ACTGGGAGCTACTGTCAGGAGCAACTCGATGAGTGCGGTCCAAACCCCTGCCAGCAGCGG 3489
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1652 SerLysValAsnLeuPhe-----CysAspProGlyPheGlnLeu 1664
Q _y		
	: : : : : : : :	3490 GCACATGTCAGTGACTTCATTGGTGGATACAGATGCGAGTGTCTCCAGGCTATCAGGGT 3549
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1665 ValGlyAsnProVal-----GlnTyrCysLeuAsnGlnGlyGln 1677
Q _y		
	: : : : : : : :	3550 GTCAACTGTCAGTATGAAGTGGATGATGCCAATCAGCCCTGCCAGAATGGAGGCACC 3609
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1678 TrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGlu 1697
Q _y		
	: : : : : : : :	3610 TGTATTGACCTTGTGAACCAT-----TTCAAAGTGCTTTGCCCAACA----- 3651
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1698 AsnGlyPheHisSerAlaAspPheTyrAlaGlySerThrValThrTyrGlnCysAsn 1717
Q _y		
	: : : : : : : :	3651 ----- 3651
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1718 AsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsn 1737
Q _y		
	: : : : : : : :	3652 -----GGCACTCGGGGCTACTCTGTGAAGAGAAC----- 3681
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1738 GlyValSerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGlu 1757
Q _y		
	: : : : : : : :	3682 -----ATTGATGACTGTGCGCGGGGTCCCCATTGCCTTAAT 3717
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1758 HisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThr 1777
Q _y		
	: : : : : : : :	3718 GGTGGTCAGTCATGCATAGGATTGGAGGCTACAGTTGTGCTGCTGTGCTGCTGTGCT 3777
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1778 GlyAspGlyLysAsnCysAlaGluProIleLys-----CysLysAla 1791
Q _y		
	: : : : : : : :	3778 GGGGAG-----CGTTGTGAGGGAGACATCAACGAGTGCCCTCTCCAACCCCTGCAGCTCT 3831
D _b		
	::: :::: :::: :::: :::: :::: :::: ::::	1792 ProGlyAsn 1794
Q _y		
	: : : : : : : :	3832 GAGGGCAGC 3840
D _b		

RESULT 4

US-08-532-384-21
Sequence 21, Application US/08532384
Patent No. 6083904
GENERAL INFORMATION:
APPLICANT: Artavanis-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
TITLE OF INVENTION: Nucleic Acids
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:

GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmuth, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 110
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 009051.5CB1
US-09-484-970B-110

Alignment Scores:
Pred. No.: 3,32e-67 Length: 763
Score: 838.00 Matches: 148
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.20% Indels: 0
DB: 4 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-484-970B-110 (1-763)

```
Qy 720 IleValIleLysGlySerProCysGluLeuProPheThrProValAenGlyAspPheIle 739
Db 763 ATTGTCAATAAAGTTCTCCCTGTGAAATCCATTACACCTGTAATGGGATTTTATA 704
Qy 740 CysThrProAspThrGlyValAenCysThrIleuThrCysLeuGluGlyTyrAspPhe 759
Db 703 TGCACCTCCAGATAATACTGGAGTCAACTGTACATTAACTTCTTGGAGGCTATGATTTC 644
Qy 760 ThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTyrLysProThr 779
Db 643 ACAGAGGGTCTACTGACAGTATTATTGTCTTATGAGATGGCTCTGGAAACCAACA 584
Qy 780 TyrThrThrGluTrpProAspCysAlaLysLysArgPheAlaAenHisGlyPheLysSer 799
Db 583 TATACCACCTGAATGGCCAGACTGTGCCAAACAAAGTTTTCACAAACACCGGTTCAATCC 524
Qy 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
Db 523 TTTGAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAAGTTTCT 464
Qy 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
Db 463 GAAGCATTTGAGACGACCTCGGGAANAATGTCCTCCATCAATTTGTAGTGTGATGAGGAC 404
Qy 840 IleAspCysArgLeuGluAenLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
Db 403 ATTGACTGCAGACTGGAGGAGAACCTGACCAAAAAATATTGCTTAGAATAATAATTATGAC 344
Qy 860 TyrGluAsnGlyPheAlaIleGly 867
Db 343 TATGAAATGGCTTTGCAATTGGT 320
```

RESULT 6

US-09-230-652-1
; Sequence 1, Application US/09230652A
; Patent No. 653775
; GENERAL INFORMATION:
; APPLICANT: Tournier-Lasserre, Elisabeth
; APPLICANT: Joutel, Anne
; APPLICANT: Bousser, Marie-Germaine
; APPLICANT: Bach, Jean-Francois
; TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATION
; FILE REFERENCE: 03715.0048-00000

CURRENT APPLICATION NUMBER: US/09/230,652A
; CURRENT FILING DATE: 1999-05-17
; EARLIER APPLICATION NUMBER: FR 96 09733
; EARLIER FILING DATE: 1996-08-01
; EARLIER APPLICATION NUMBER: FR 97 04680
; EARLIER FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: PCT/FR97/01433
; EARLIER FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(7041)
; OTHER INFORMATION: human ADNC No. 6537775sch 3
US-09-230-652-1

Alignment Scores:
Pred. No.: 1.55e-64 Length: 8091
Score: 829.00 Matches: 259
Percent Similarity: 37.07% Conservative: 85
Best Local Similarity: 27.91% Mismatches: 287
Query Match: 8.11% Indels: 298
DB: 4 Gaps: 46

US-09-977-053-6 (1-1842) x US-09-230-652-1 (1-8091)

```
Qy 988 CysArgProGlySerValLeuArg---GlyArgMetCysValAen----- 1001
Db 562 TGCCGGGTGGGTGAGCCCTGCCGCCATGTTGGCCACCTGCTCCACACACCTGGCTCTTC 621
Qy 1002 -----CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer--- 1017
Db 622 CGCTGCCAGTGTCCAGCTGGCTACACAGGGCCACTA-----TGAGAGAACCCC 669
Qy 1018 -----CysArgIleGlySerTyrGlnAspGluGluGlyGln 1029
Db 670 GCGGTGCCCTGTGGCCCTCACCATGCGTAACGGGGGCACCTGCAGGCAGAGTGGCGAC 729
Qy 1030 Leu-----GluCysLysLeuCysPro-----SerGlyMetTyrThrGluTyrIleHis 1045
Db 730 CTCACCTACGACTGTGCTGCTCTCTCTGAGGGTTGAGGGTCAGAAATTGTGAAGTG--- 783
Qy 1046 SerArgAsnIleSerAspCysLysAlaGln-----CysLysGln 1058
Db 784 -----AACGTGAGCGACTGTGCCAGSACCGCATGTCTCAATGGGGGACATGCGTGGAT 837
Qy 1059 GlyThrTyrSerTyrSer-----GlyLeuGlu 1067
Db 838 GGCTCAACACCTATAACTGCCAGTGCCTCTGAGTGGACAGGCCAGTTCGACGGAG 897
Qy 1068 ThrCysGluSerCysProLeu-----GlyThrTyrGlnPro 1079
Db 898 GACGTGATGAGTGTGAGCTGCAGCCCAACGCTGCCACAAATGGGGTACCTGCTTCAAC 957
Qy 1080 LysPheGlySerArgSerCysLeu-----SerCysProGlu 1091
Db 958 ACGCTGGTGGCCACAGCTGCTGTGTGTCAATGGCTGGACAGGTGAGAGCTGCAGTCAG 1017
Qy 1092 Asn-----ThrSerThrValLysArgGlyAlaVal-----AsnIle 1103
Db 1018 AATATCATGACTGTGCCACAGCGTGTCTTCCATGGGGCCACTGCGCATCAGCGGTG 1077
Qy 1104 SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys 1123
Db 1078 GCTTCTTTTACTGTGCTGCCCTGCCCATGGCAAG-----ACTGGCTCTCTG---TGT 1125
Qy 1124 His-----ProCysProArgAspTyrTyr-----Gln 1132
Db 1126 CACTGGATGACCGCTGTGTGTGTCAGCAACCCCTGCCACGAGGATGCTATCTGTGACACAA 1185
```



```

QY 1615 lleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArgLeu 1634
DB 2726 TTCCAGGGGCTGCACTGTGAGGAGAAGACTAACCCAGCTGTGTCAGACAGCCCTGCAGG 2785
QY 1635 GlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerIleVal 1654
DB 2786 AAC-----AAGGCACTGCCAAGACACACCTCGAGGGGCGCG-----2824
QY 1655 AsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnIleCysLeuAsn 1674
DB 2825 ---TGCCCTCTGAGCCCTGGCTAT-----ACAGGAAGCAGCTGCCAGACTCTGTATGAC 2875
QY 1675 GlnGlyGlnIleThrGlnProLeuProHisCysGluArgIleSerCysGlyValProPro 1694
DB 2876 TTG-----TGTCGCCGGAAGCCCTGTCCA-----2899
QY 1695 ProLeuGluAsnGlyPheHisSerAlaAspPheTyrAlaGlySerThrValThrTyr 1714
DB 2900 -----CACACTGCTCGATGCTCCAGAGTGGGCGCCCTCGTTCCAGTGC 2941
QY 1715 GlnCysAsnAsnGlyTyrTyr---LeuLeuGlyAspSerArgMetPheCysThrAspAsn 1733
DB 2942 CTGTGCTCCAGGATGAGCAGGGGCTCTGTGACTTCCACTGCTCCAGAGGCC 3001
QY 1734 GlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaValGlySer 1753
DB 3002 CGCATGAGCCAGGACATAGATCTCTGCGCTG-----3034
QY 1754 AspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysVal 1773
DB 3035 ---TGCCAGATGAGGAGCTCTGTATTGACACGGGCTCTCTCTATTCTTCGCGCTCCCT 3091
QY 1774 ProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysCysLysAlaProGly 1793
DB 3092 CTGGAATTC-----CAAGCAAGTTATGCCAGGATAATGTGAAC-----CCCTGC 3136
QY 1794 AsnProGluAsnGlyHisSerSerGlyGluIleTyrThrValGlyAlaAlaValThrPhe 1813
DB 3137 GAGCCCAATCCCTGCCATCAGGGCTCTACCTGTGTGCTCCAGCCAGTGGCTATGTCTGC 3196
QY 1814 SerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle-----ThrCysLeuGlu 1831
DB 3197 CAGTGTGCCCCAGGCTATAGGGAGACAGAACTGCTCAAAAGTACTTGCAGCTTGTGAGTCC 3256
QY 1832 SerGlyGluTrpAsnHis 1837
DB 3257 CAGCCCTGCCACACACCAC 3274

```

RESULT 9

US-09-467-997-6
Sequence 6, Application US/09467997
Patent No. 6379925

GENERAL INFORMATION:

APPLICANT: Kitajewski, Jan
APPLICANT: Uytendaele, Hendrik
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
FILE REFERENCE: 53863-A-PCT-US
CURRENT APPLICATION NUMBER: US/09/467,997
CURRENT FILING DATE: 1999-12-20
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 6677
TYPE: DNA
ORGANISM: mouse

US-09-467-997-6

Alignment Scores:
Pred. No.: 9,086-59
Score: 764.50
Percent Similarity: 36.35%
Best Local Similarity: 26.05%

Length: 6677
Matches: 273
Conservative: 108
Mismatch: 399

```

Query Match: 7.48% Indels: 268
DB: 4 Gaps: 47
US-09-977-053-6 (1-1842) x US-09-467-997-6 (1-6677)
QY 927 ProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGlnArgLeuLeuGlnThrLeu 946
DB 530 CTGACCCCTGAGGAGTACCCCACTCTCCAGAGATGGTGGCAGCTGCCAAGCCCTGCTC 589
QY 947 GluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGln 966
DB 590 CCCACACCCCAAGCTCCCGTAGTCTACTTCTCCACTGACCCCTCACTTCTCTCGCACC 649
QY 967 LeuAlaSerGluIleLeuLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986
DB 650 TGCCCTCTGCGC---TTACCCCGTGTATGATGCAAAACCATCTGGAAGAGCTCTGTCCA 706
QY 987 -----PheCysArgProGlySer-----ValLeuArgGlyArgMetCysVal 1000
DB 707 CTTCTTTCTGTCCACCGGGGTCACTGCTATGTTGAGGCTCAGGCGCCGCCACAGTGC 766
QY 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
DB 767 TCTGCGAGCTGGG-----TGACAGGTGAGCAATGCCAGCTC 805
QY 1021 GlySerTyrGluAspGluGlu-----Gly 1028
DB 806 CGAGACTTCTGTGCTAGCCCAACCCCTGTGCCAACCGAGGGGTGTGCTGGCCACATACCC 865
QY 1029 GlnLeuGluCysValLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAsn 1048
DB 866 CAGATCCAGTGGCGC---TGTCCACTGGG---TTGAGGCTCACACCTGTGAAAGCGAC 919
QY 1049 IleSerAspCys-----LysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065
DB 920 ATCAACAGAGTGTCTTCTGCGAGCGGACCCCTGCGCCCTCAGGGCACCTCTGCCATAAC 979
QY 1066 LeuGluThrCysGluSer---CysProLeuGlyThrTyrGlnProLysPheGlySerArg 1084
DB 980 TTGGGTCTCTACCACTGTCTCTGCGCTGTGGGGCAGGAAGGTCCCGAGTCAAGCTCAGG 1039
QY 1085 -----SerCysLeuSer-----CysProGlu 1091
DB 1040 AAGGAGCTGCGCTCTCTGAGAGCTGTCTCAATGGGGGACCTGCCAGCTGGTCCAGAG 1099
QY 1092 AsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValPro----- 1109
DB 1100 GGACACTCCACC-----TTTCATCTCTGCTCTGTCTGCCCCAGGT 1138
QY 1110 -----CysProGlu 1112
DB 1139 TTACCGGGCTGGACTGTGAGATGAACCCAGATGACTGTGTGAGGACCAAGTGTCAAGAC 1198
QY 1113 GlyLysPheSerArgSerGlyLeu-----MetProCysHisProCysProArgAspTyr 1130
DB 1199 GGGGCCACCTGTCTGATGGCTGGATACCTACACTGC---CCCTGCCCCCAAGACATGG 1255
QY 1130 ----- 1130
DB 1256 AAGGGCTGGAGCTGCTCTGAAGATATAGATGAATGTGAAGCCCGGGTCCCTCGCTGC 1315
QY 1131 -----TyrGlnProAsnAlaGlyAlaPheCysLeuAlaCysProPhe 1145
DB 1316 AGGAACGGTGGACCTGCCAGAACACAGCTGGCAGCTTTCACCTGTGTGCTGGTGGTGGC 1375
QY 1146 TyrGly-----ThrThrProPheAla 1152
DB 1376 TGGGAGGTGCGAGGCTGTGAGGAGAACCTGATGATGCTGAGTGCAGCTGCACCTGTGCCCG 1435
QY 1153 GlySerArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGlu 1172
DB 1436 GGATCCACCTGCATCGACCGTGTGGGCTCTTTCTCC-----1471

```

1173	Qy	SerValValProProAlaSerLeuGlyHisIleLysLysArgHisGluLleSerSerGln	1192
		:::	
1472	Db	TGCCTGCCCCCACTGGACGCACACAGGCTCTCTGCGCACCTGGAAAGCATG-	1522
		:::	
1193	Qy	ValPheHisGluCysPhePheAenProCysHisAasnSerGlyThrCys-	1210
		:::	
1523	Db	-----TGTTTGAGTCACCGCTGCACCTGAATGCCAGTCAGCAGCAACCTGTGCACCA	1570
		:::	
1211	Qy	LeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGlyLeuHisCysGluThr	1230
		:::	
1571	Db	CTGACAGGCTCCACCTCTGCATATGCAGCTGGCTACTCAGATCCACTGTGCACCA	1630
		:::	
1231	Qy	AspIleAspGluCys-----SerProLeuProCysLeuAasnAenGlyVal	1245
		:::	
1631	Db	GATCTGATGAGTGCCAAATGGCCCCACAGGACCCAGTCCCTCGCAACATGGGGGTCC	1690
		:::	
1246	Qy	CysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGlnArg	1265
		:::	
1691	Db	TGCATCAACACCCCTGGCTCTTAAGTCTTGCTTGCTGGCTGTGTACAGGGCTCCCGC	1750
		:::	
1266	Qy	CysGluGluAasnIleAasnGluCysSerSerProCysLeuAasnLysGlyIleCysVal	1285
		:::	
1751	Db	TGTGAAGCTGACCAACATGAGTGCTGTACAGCCCTGCCACCCAGGACGACCTGCCTG	1810
		:::	
1286	Qy	AspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHisCysGlu	1305
		:::	
1811	Db	GACCTGCTTGCAGGTTCACACTGCTCTGCTGCTGCTGCTGTGTACAGGGTCTCTGTGAG	1870
		:::	
1306	Qy	ThrGluValAasnGluCysGlnSerAsnProCysLeuAasnAlaValCysGluAspGln	1325
		:::	
1871	Db	GTGAGGTCAATGAGTGACCTCTAATCCCTGCTGTGAACCAAGCTGCTGCCATGACCTG	1930
		:::	
1326	Qy	ValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGlyLysAasn	1345
		:::	
1931	Db	CTCAACGGCTTCAGTGCTCTGCTCTTCTGATTACCGGGGCCCGCTGCCGAGACCGCTGGA	1990
		:::	
1346	Qy	ValAspGluCysLeuSerGlnProCysLysAasnGlyAlaThrCysLysAspGlyAlaAasn	1365
		:::	
1991	Db	ATGACCGAGTGTAGCACACCCCTGTGTCGAATGGGGGGCGCTGCCGAGACCGCTGGA	2050
		:::	
1366	Qy	SerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAasnIleAasn	1385
		:::	
2051	Db	GCCTTCTACTGCGAGTGTCTCCAGGCTTTGAAGGGCCACACTGTGTGAAGAAGTGGAC	2110
		:::	
1386	Qy	GluCysGlnSerAenProCysArgAasnGlnAlaThrCysValAspGluLeuAasnSerTyr	1405
		:::	
2111	Db	GAATGCTGAGTGACCCCTGTGTCGTGGAGCCAGCTGTCTTGAATCTCCCGGAGCATTC	2170
		:::	
1406	Qy	SerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSerThr---	1424
		:::	
2171	Db	TTCTGTCTCTGCGCTCTGTGTTTACAGGTCAACTTTGTGAGGTTCCTTTGTGACCCCC	2230
		:::	
1425	Qy	-----GlyPheAasnLeuAspPheGlu-----	1431
		:::	
2231	Db	AACATGTGCCAACTGTGACAGCAATGCCAGGTACAGGAACACAGAGCCCCCTGCTCTGC	2290
		:::	
1432	Qy	ValSerGlyIleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHis-AlaLe	1451
		:::	
2291	Db	CCTGACGGAGTCTGCTGTGTCTTCCTCCGAGGACAACTGCCCTCTCAGCATGGCCAT	2350
		:::	
1451	Qy	uThr-----CysThrPheTrpMetLysSerSerAspAspMetAasnTyrGlyThrPr	1468
		:::	
2351	Db	TGCCAGAGATCTTGTGTGTGTGATGAGGGCTGGACTGGACCAAGAT	2399
		:::	
1468	Qy	oIleSerTyrAlaValAspAasnGlySerAasnThrLeuLeuLeuThrAspTyrAasnG	1488
		:::	
2400	Db	-----GCCGACAGCAGAAC-----	2411
		:::	
1488	Qy	yTrpVal-LeuTyrValAasnGlyArgGluLysIleThrAasnCysProSerValAasnAspG	1508
		:::	
2412	Db	TGGGTGGCTGCATCTCCAC-----ACCTGTGCCCA-----	2442
		:::	
1508	Qy	lyArgTrpHisHisIleAlaIleThrTrpTrpSerAlaAasnGlyIleTrpTrpLysValT	1527
		:::	

Db	2443	 -----TGGGGGGACCTGGCCACCCACA-----	 -----GCATCTGCTCAACTGT	2482
Qy	1527	yrileAspGlyIysLeuSerAsp-----	GlyGlyAlaGlyLeuSerVal-	1541
Db	2483	ACCTGCCCTCGCAGGTACATCGGGGTGACCTGTAGTCAGAGAGGTGACAGCTTGTCACTCA	2542	
Qy	1542	GlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluInAspLysLysGly	1561	
Db	2543	GGGGCCCTGTCTCAATGGTGGCTCTCGCAGCATC-----	CGACCT	2581
Qy	1562	GlucIyPheSerProAla-----	GlusPheValGlySerIleSerGlnLeuAsnLeu	1579
Db	2582	GAGGGCTATTCTTCGCACCTTCCTCCCAAGTCACACAGGTGCCACTGCCAGACTGCCGTG	2641	
Qy	1580	TriAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu	1599	
Db	2642	---GACCACACTGTGTGTCT-----	GCCTGTGC-----	2665
Qy	1600	LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeu-----	SerGly	1614
Db	2666	CTCAATGGGGGTACCTGTGTGAACAGACCTGGCATTCTTCTGTGCCTCTGTGCCACTGGC	2725	
Qy	1615	IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArgLeu	1634	
Db	2726	TTCAGGGGCTGCACCTGTGAGGAGAGACATAACCCAGCTGTCCAGACACCCCTCGCAGG	2785	
Qy	1635	GlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLysVal	1654	
Db	2786	AAC-----	NAGGCACCTGCCAGACACACCTCGAGGGGCCCGC---	2824
Qy	1655	AsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsn	1674	
Db	2825	---TGCCTCTGCAGCCCTGGCTAT-----	ACAGGAGCAGCTGCCAGACTCTCGATAGAC	2875
Qy	1675	GlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProPro	1694	
Db	2876	TTG-----	TGTCCCGGAGCCCTGTCCA-----	2899
Qy	1695	ProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyr	1714	
Db	2900	-----	CACACTGCTCGATGCTCCAGAGTGGCCCTCGTTCAGTGC	2941
Qy	1715	GlnCysAsnAsnGlyTyrTyr-----	LeuLeuGlyAspSerArgMetPheCysThrAspAsn	1733
Db	2942	CTGTGCTCTCCAGGATGGACAGGGGCTCTCTGTGACTTTCCTCCACTGTCTGCCAGAGGCC	3001	
Qy	1734	GlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaValGlySer	1753	
Db	3002	GCGATGAGCCAGGACATAGAGATCTCTGGCCTG-----	3034	
Qy	1754	AspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysVal	1773	
Db	3035	---TGCCAGATGGAGGCCCTCTGTATTGACACGGGCTCTCTCTATTTCGCGCTGCCCT	3091	
Qy	1774	ProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysCysLysAlaProGly	1793	
Db	3092	CCTGGATTC-----	CAAGGCAAGTTATGCCAGGATAATGTGAAC-----	3136
Qy	1794	AsnProGluAsnGlyHisSerSerGlyLulIeTyrThrValGlyAlaAlaValThrPhe	1813	
Db	3137	GAGCCCAATCCCTGCCATACACGGGTCTACTGTGTGCTCAGGCCAGTGGCTATGTCTGC	3196	
Qy	1814	SerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle-----	ThrCysLeuGlu	1831
Db	3197	CAGTGTCCCCAGGCTATGAGGACAGACAACTGCTCAAAGTACTTTCAGCTTGTTCAGTCC	3256	
Qy	1832	SerGlyGluTrpAsnHis	1837	
Db	3257	CAGCCCTGCCACACACAC	3274	

; Sequence 6, Application US/09214278

; Patent No. 6291210

; GENERAL INFORMATION:

; APPLICANT: Sakano, Seiji

; APPLICANT: Itoh, Akira

; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE

; FILE REFERENCE: KP-8576

; CURRENT APPLICATION NUMBER: US/09/214,278

; CURRENT FILING DATE: 1999-01-26

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 4208

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (409)..(4062)

; NAME/KEY: sig_peptide

; LOCATION: (409)..(501)

; NAME/KEY: mat_peptide

; LOCATION: (502)..(4062)

; US-09-214-278-6

Alignment Scores:

Pred. No.:	1,05e-57	Length:	4208
Score:	749.00	Matches:	257
Percent Similarity:	35.13%	Conservative:	127
Best Local Similarity:	23.51%	Mismatches:	347
Query Match:	7.33%	Indels:	362
DB:	3	Gaps:	48

US-09-977-053-6 (1-1842) x US-09-214-278-6 (1-4208)

QY	834	CysSerAspAlaGlu	-----AspIleAspCysArgLeuGluGluAsnLeuThrIys	850
DB	571	TGCGGGCGCCCGGACCCGAGGACCCGAGTGCACCGGAGCTGCACATAC	-----	630
QY	851	---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe	-----AlaIleGlyPro	868
DB	631	TTCAAAAGTGTGCTCAAG	-----GAGTATCAGTCCCGGTCAACCTCAAGCCGCGGGGCC	681
QY	869	GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspPheLeuAspThrVal	888	
DB	682	TCAGCTTCGGCTCAGGG	-----	699
QY	889	GlnGluThrAlaThr-SerIleGlyAsnAlaLysSerArgIleLysArgSerAlaPro	908	
DB	700	-----TCCAGCCTGTATCGGGGCAACACCTTCAACCTCAAGCCAGCCCGGC	750	
QY	909	LeuSerAspTyrLysIleLysLeuIlePheAsnIle	-----ThrAla	922
DB	751	AACGACCGCAACCGCATCGTGTCTTCAGTTTCAGTTCGCGCCGAGGTCTTATACGTTG	810	
QY	923	SerValProLeuProAspGluArgAsnAspThrLeuGluTyrGlu	-----	937
DB	811	CTTGTGGAGCGTGGGATTCAGTATGACCCGTTTCAACCTGACAGTATATTGAAAG	870	
QY	938	-----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr	950	
DB	871	GCTTCTCACTCGGGCATGATACCCCGCGGAGTGGCAGACGCTGAAGCAGACACG	930	
QY	951	AsnLysLeuLys	-----ArgThrLeuAsnLysAspProMetTyrSerPhe	965
DB	931	GGCGTTGGCCACTTTGAGTATCAGATCCGCGTGCAGCTGTGATGACTACTACTATGCTTT	990	
QY	966	GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrIysLysAlaSer	985	
DB	991	-----GGCTGCAAT	999	
QY	996	ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly	1005	
DB	1000	AAGTTCTGCGGCGCCAGAGATGACTTCTTTGA	-----	1032

QY	1006	ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp	1025	
DB	1033	-----CACTATGCTGTGAC	1068	
QY	1026	GluGluGlyGlnLeu	-----LeuCysProSerGlyMetTyrThr	1041
DB	1069	ATGGAAGGTGGATGGCCCGGATGTAACAGAGACTATTTCGCGACAGAGCTGCAGTCT	1128	
QY	1042	GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr	1061	
DB	1129	AAGCATGGGTCTTGCAAACTCCAGGTGACTGCGGTGCGAG	1173	
QY	1062	SerTyrSerGlyLeuGluThrCysGluSerCys	-----ProLeuGlyThrTyr	1077
DB	1174	GGCTGGCAAGGCTG	-----TACTGTGATAAGTGCATCCACACCCGGGATCGTCCACGGC	1230
QY	1078	-----GlnProLysPheGlySerArgSerCysLeu	1087	
DB	1231	ATCTGTATGAGCCCTGGCAGTGCCTCTGTGAGACCACTGGGGCGCCAGCTCTGT	1287	
QY	1088	SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly	1107	
DB	1288	-----GACAAAGATCTCAATTACTGTGGG	1311	
QY	1108	Val	-----ProCysProGluGly	1116
DB	1312	ACTCATCAGCCCTGTCTCAACGGGGGAACCTTTAGCAACACAGCCCTGCACAAATATCAG	1371	
QY	1117	ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly	1136	
DB	1372	TGTTCC	-----TGCCCTGAGGGGTATTCAGGACCCCACTCT	1407
QY	1137	LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer	1156	
DB	1407	-----	1407	
QY	1157	IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluSerValPro	1176	
DB	1407	-----	1407	
QY	1177	ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu	1196	
DB	1408	-----GAATTGCTGAG	1425	
QY	1197	CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal	1216	
DB	1426	TGCCTCTCTGATCCCTGTCAACAGAGGAGCAGCTGTAAAGGAGACCTCCCTGGGCTTTGAG	1485	
QY	1217	CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer	1236	
DB	1486	TGTGAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTTGATGCTTCT	1545	
QY	1237	ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu	1256	
DB	1546	CTAATAACTGTGTCCACCGGGGACCTCCAGGACTTGTAAAGATTTAAAGTGTG	1605	
QY	1257	CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer	1276	
DB	1606	TGCCCCCAGACAGTGGGAAACCTGCCAGTTAGATGCAATGAATGTGAGGCCAAA	1665	
QY	1277	ProCysLeuAsnLysGlyIleCysValaspGlyValAlaGlyTyrArgCysThrCysVal	1296	
DB	1666	CCTTGTGTAACGCCCAATCTCTGTAAGAAATCTCATTCGCCAGCTACTACTCGGACTGCTT	1725	
QY	1297	LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys	1316	
DB	1726	CCCGCTGATGGCTCAGAAATTTGTACATAAATAATTAATGACTGCTTGGCCAG	1782	
QY	1317	LeuAsnAlaValCysGluAspGlnValGlyPheLeuCysLysCysProProGly	1336	
DB	1783	CAGAATGACGCTCTCTGCGGATTTGGTTAATGGTTATCGCTGTATCTGTCCACCTGGC	1842	

1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
1843 TATCCAGCGCATCACTGTGAGAGAGACATCGATGAATGTGCCAACCCCTGTTGAAT 1902
1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
1903 GGGGGTCACTGTTCAGATGAATCAACAGATTCAGTGTCTGTCTCCACTGGTTCCTCT 1962
1377 GlySerHisGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGluAla 1396
1963 GGAACCTCTGTGAGCTGAGACATCATATTGTGAGCTTAATCTTCCAGACAGCGTGCC 2022
1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
2023 CAGTGTACAAACCGTGCAGTCACTATTCTTCAAGTGCCTGAGGAGCTATGAGGCAAG 2082
1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
2083 AACTGTCTCACACCTGAAGAGCACTGCGCGCAGCACCCCTGTGAAGTG----- 2130
1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
2131 -----ATGAC-----AGCTGCACAGTGCC 2151
1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
2152 ATGCTTCCAACGAC-----ACACTGAAGGGGTGGGTATATTTCCTCC 2196
1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
2197 AAC-----GTCTGTGTCTCTCAC 2214
1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAla 1514
2215 GGGGAG-----TGCAAGAGTCAGTCGGAGGCAATTC----- 2247
1515 IleThrTyrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
2248 ---ACCTGTGACTGTAAACAAGGCTTCACGGGAACATACTGCCATGAATAATTAATGAC 2304
1535 GlyGlyAlaGlyLeuSerValGlyLeuProGlyLeuProGlyGlyAlaLeuValLeuGly 1554
2305 TGTGAGAGCAAC-----CCTGTAGAAACGGTGGCCTTGCATCGATGGT 2349
1555 GlnGluGlnAspLysGlyGluGlyPheSerProAlaGlnSerPheValGlySerIle 1574
2350 GTCACTCTCTACAG-----TGCACTGTGTAGTCGCGCTGGAGGGGCGCTAC 2397
1575 SerGlnLeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
2398 TGTGAACCAATATTATGACTGTCAGCCAGCAGAACCC----- 2433
1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
2434 ---TGCCACAAT----- 2442
1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
2443 -----GGGGCAGCTGTGCGACATGTGTCAATGACTTCTACTGTACTGTAAAAATGGG 2496
1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
2497 TGGAAAGGAAGACCTGCCACTCAGTGCAGCTCAGTGTGTGATGAGGCGCAGCTGCAACAC 2556
1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
2557 GGTGGCACCTGCTATGATGAGGGGGATGCTTTTAAG----- 2592
1672 CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
2593 TGCATGTCTCTGGCGGTGGGAAGGAGCAACCTGTGTAACTAGCCCGAAGACAGTAGCTGC 2652
1692 ValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThr 1711

2653 CTGCCCAACCCCTGCCAATATGG----- 2676
1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
2677 -----GGCACATGTGTGTGTCACGGGAGTCTCTTACGTGCGTC 2715
1732 AspAsnGlySerTyrAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
2716 TGCAAGGAAGGCTGGAGGG-----CCCATCTGTGTCTCAGATACC----- 2757
1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
2758 ---AATGACTGACGCTCATCTCTTACACAGCGCACCTGTGTGGATGAGATCAATGGCTAC 2814
1767 SerTyrIleCysSerCysValProTyrThrGlyAsp----- 1779
2815 TGGTACCGTTCGATGTGCCCCGGGTTTGTGGCCGAGTGCAGATAAATCAATCAAT 2874
1780 -----GlyLysAsnCysAlaGluProIle----- 1787
2875 GAATGCCAGTCTTCACTTGTGCTTGTGGAGCGACCTGTGTGGATGAGATCAATGGCTAC 2934
1788 LysCysLysAlaProGlyAsnProGlyAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
2935 CGGTGTGTCTGCTCT-----CCAGGGCACAGTGGTGCCAAAG----- 2970
1808 GlyAlaAlaValPheSerCysGlnGlyTyrGlnLeuMetGlyValThrLysIle 1827
2971 -----TGCCAGGA-----GTTTCAGGAGGA 2991
1828 ThrCysLeuGluSerGlyTyrAsnHisLeuIlePro 1840
2992 CTTTGCATCACATGGGAGT-----GTGATACCA 3021

RESULT 11

US-09-855-722-6
; Sequence 6, Application US/09855722
; Patent No. 6638741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; NAME/KEY: sig_peptide
; LOCATION: (409)..(501)
; NAME/KEY: mat_peptide
; LOCATION: (502)..(4062)
US-09-855-722-6

Alignment Scores:

Pred. No.:	1,05e-57	Length:	4208
Score:	749.00	Matches:	257
Percent Similarity:	35.13%	Conservative:	127
Best Local Similarity:	23.51%	Mismatches:	347
Query Match:	7.33%	Indels:	362
DB:	4	Gaps:	48

US-09-977-053-6 (1-1842) x US-09-855-722-6 (1-4208)

QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
DB 571 TCGCGCGCGCCCGAACCACCGGAGACCAAGTGCACCGGACGAGTCTGACACATAC 630
QY 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
DB 631 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGGTCAACGCGCGGGGGCC 681
QY 869 GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
DB 682 TGCAGCTTCGGCTCAGG----- 699
QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
DB 700 -----TCCAGCGCTGATCGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGC 750
QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
DB 751 AACGACCGCAACCGCATCGTCTGCTTTCAGTTTCGCTGCGCGAGGTCTATACGTTG 810
QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluThrIleThr 950
DB 811 CTTGTGGAGCGTGGGATTCAGTAATGACACCGTTCAACCTGCACAGTATTATTGAAAG 870
QY 938 -----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
DB 871 GCTTCTCACTCGGCGATGATCAACCCAGCGGAGTGGCAGCGCTGAAGCAGAACACG 930
QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
DB 931 GCGGTGCCCACCTTTCAGTATCAGATCCCGGTGACCTGTGATGACTACTACTATGCTTT 990
QY 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
DB 991 ----- 999
QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
DB 1000 AAGTTTCGCGCGCCAGAGATGACTCTTTTGA----- 1032
QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
DB 1033 -----CACTATGCTGTGAC-----CAGAAATGGCAACAAACTTGC 1068
QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
DB 1069 ATGGAAGCTGGATGGGCGCCGAATGTAACAGAGACTATTTCGCGACAGGCTGCAGTCT 1128
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 1129 AAGCATGGGTCTTGCAAACTCCAGAGTGAAGTGCAGTGCAG-----TAC 1173
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
DB 1174 GGTGCGCAAGGCTG---TACTGTGATAAGTGCATCCACACCGGATGCGTCCACGCG 1230
QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
DB 1231 ATCTGTAATGAGCCCTGCGAGTGCCTCTGTGAGACCAACTGGGGCGCGGCTCTGT--- 1287
QY 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
DB 1288 -----GACAAAGATCTCAATTACTGTGG 1311
QY 1108 Val-----ProCysProGluGly-----LysPheSer 1116
DB 1312 ACTCATCAGCGGTCTCAACGGGGGAACTTTAGCAACACAGGCCCTGACAAATATCAG 1371
QY 1117 ArgSerGlyLeuMetProCysHisProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
DB 1372 TGTTC-----TGCCTGAGGGGTATTACAGACCCCAACTCT--- 1407
QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSer 1156

DB 1407 ----- 1407
QY 1157 IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValValPro 1176
DB 1407 ----- 1407
QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
DB 1408 -----GAATTTGCTCAG-----CAGGCC 1425
QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnIleLeuGlyArgGlyTyrVal 1216
DB 1426 TGCCTCTCTGATCCCTGTCAACAGAGCAGCTGTAAGGAGACCTCCCTGGGCTTTGAG 1485
QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
DB 1486 TGTGAGTGTTCACAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTCT 1545
QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
DB 1546 CTTAATACTGTTCCACCGGGGACCTGCCAGACTTGGTTAACGGATTTAAGTGTGTG 1605
QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer 1276
DB 1606 TGCCTCCACAGTGGACTGGAAAACGTGCCAGTTAGATGCAATGAATGTGAGGCCAAA 1665
QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
DB 1666 CTTGTGTAAACGCCAAATCCTGTAGAATCTCATTCGCCAGCTACTCTCGACTGTCTT 1725
QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
DB 1726 CCGGCTGGATGGTGCAGATTGTGATAAATATTAAAGTACTGCTTGGCCAG---TGT 1782
QY 1317 LeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGly 1336
DB 1783 CAGAAATGACGCTCTCTGCTGGATTGTTGTTAATGTTATCGCTGTATCTGTCCACCTGGC 1842
QY 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
DB 1843 TATGCAAGCGATCCTCTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTGAA 1902
QY 1357 GlyValaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
DB 1903 GGGGTCACTGTGAGAAATCAACAGATCCAGTGTCTGTGCCACTGCTTCTCT 1962
QY 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
DB 1963 GGAACCTCTGTGAGTGGACATCGATTATTGTAGACCTAATCCCTGCCAGAACGGTGC 2022
QY 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
DB 2023 CAGTGTACACCGTGCAGTACTATTCTGCAAGTGCCTCCAGGACTATGAGGGCAG 2082
QY 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
DB 2083 AACTGCTCACCTGAAGACCACTCCCGCAGCACGCCCTCTGTGAGTG----- 2130
QY 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
DB 2131 -----ATTGAC-----AGCTGCACAGTGGCC 2151
QY 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
DB 2152 ATGGCTTCCAAAGAC-----ACACCTGAAGGGGTGGGTATATTTCTCTCC 2196
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
DB 2197 AAC-----GTCTGTGGTCTCTCAC 2214
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAla 1514

Db 2215 GCGAAG-----TGCACAGTCAGTCGGGAGGCAATTC----- 2247
Qy 1515 IleThrTrpThrSerAlaAsnGlyIleThrIleAspGlyLeuSerAsp 1534
Db 2248 ---ACCTGTGACTGTAACAAGGCTTCACGGGAACATACCTCCATGAATAATTATTAATGAC 2304
Qy 1535 GlyGlyAlaGlyLeuSerValGlyLeuProGlyGlyGlyAlaLeuValLeuGly 1554
Db 2305 TGTGAGAGCAAC-----CCTGTGAGAACGGTGGGCACTTGTGATCGATGCT 2349
Qy 1555 GlnGlnGlnAspIleValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1574
Db 2350 GTCAACTCCTTACAAG-----TGATCTGTAGTACGCTGGGAGGGGCGCTTAC 2397
Qy 1575 SerGlnLeuAsnLeuThrAspTyrValLeuSerProGlnGlnValLeuSerLeuAlaThr 1594
Db 2398 TGTGAACCAATATTAATGACTGACCGACAGAACCC----- 2433
Qy 1595 SerCysProGluGluLeuSerGlyGlyAsnValLeuAlaTrpProAspPheLeuSerGly 1614
Db 2434 ---TGCCACAAAT----- 2442
Qy 1615 IleValGlyIleValIleAspSerIlePheCysSerAspCysProArg--- 1633
Db 2443 ---GGGGCACGTGGCGGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGGG 2496
Qy 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuIleProGlySerIle 1653
Db 2497 TGGAAAGGAAGACCTCCACTCAGTCAGTGTGATGAGGCGCACCTGCAACAAAC 2556
Qy 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnIle 1671
Db 2557 GGTGGCACCTCTATGATGAGGGGATCTTTTAAG----- 2592
Qy 1672 CysLeuAsnGlnGlyIleThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
Db 2593 TGCATGTGTCTGGCGGCTGGGAAGGAACAACCTGTAAACATAGCCCGAACAAGTAGCTGC 2652
Qy 1692 ValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThr 1711
Db 2653 CTGCCCCAACCCCTGCCATAATGG----- 2676
Qy 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
Db 2677 -----GGCACATGTGTGTCAACGGGAGTCTTTTACGTGGGTC 2715
Qy 1732 AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
Db 2716 TGCAGGAAGGCTGGGAGGG-----CCATCTGTGTCTCAGATACC----- 2757
Qy 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
Db 2758 ---AATGACTGCAGCCCTCATCCCTGTATCAACAGCGGACCTGTGTGGATGGAGACAAC 2814
Qy 1767 SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779
Db 2815 TGTACCGGTGGAGATGTGCCCCGGTTTCTGGGCGGAGTGCAGATAAATCAATCAAT 2874
Qy 1780 -----GlyIleAsnCysAlaGluProIle----- 1787
Db 2875 GAATGCGCAGTCTTCACTTGTGCGCTTGGAGCGACCTGTGTGGATGAGATCAATGGCTAC 2934
Qy 1788 LysCysAlaValProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
Db 2935 CGGTGTCTGCCCT-----CCAGGGCACAGTGTGTGCAAG----- 2970
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
Db 2971 -----TGCCAGGA-----GTTTCAGGAGGA 2991
Qy 1828 ThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 2992 CCTTGCATCACCATGGGGAGT-----GTGATACCA 3021

RESULT 12

US-09-199-865-2
; Sequence 2, Application US/09199865
; Patent No. 6433138
; GENERAL INFORMATION:
; APPLICANT: Zimrin, Ann B.
; APPLICANT: Maciag, Thomas
; APPLICANT: Wong, Michael K.K.
; APPLICANT: Pepper, Michael S.
; APPLICANT: Montesano, Roberto
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS
; FILE REFERENCE: 0036-101
; CURRENT APPLICATION NUMBER: US/09/199,865
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: 60/018,841
; EARLIER FILING DATE: 1996-05-31
; EARLIER APPLICATION NUMBER: PCT/US97/09407
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-199-865-2

Alignment Scores:
Pred. No.: 3,89e-57 Length: 5458
Score: 745.00 Matches: 257
Percent Similarity: 34.95% Conservative: 125
Best Local Similarity: 23.51% Mismatches: 349
Query Match: 7.23% Indels: 362
DB: 4 Gaps: 48

US-09-977-053-6 (1-1842) x US-09-199-865-2 (1-5458)

Qy 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrIys 850
Db 133 TGCAGCGCGCGCGCGGAGACCGCGGAGACCGCAAGTGCACCCCGGACGATGTGACACATAC 192
Qy 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
Db 193 TTCAAGTGTGCTCAG-----GAGTATCAGTCCCGGTCCACCGCGGGGGGCC 243
Qy 869 GlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
Db 244 TGCAGCTTGGGCTCAGG----- 261
Qy 889 GlnIleThrAlaThrSerIleGlyAsnAlaIleSerSerArgIleLysArgSerAlaPro 908
Db 262 -----TCCAGCGCTGTATCAGCGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGC 312
Qy 909 LeuSerAspTyrIleLysLeuIlePheAsnIle-----ThrAla 922
Db 313 AACGACCGGAACCGCATGTGCTCTTTCAGTTTGGCTGGCGGAGGTCTCTATAGTTG 372
Qy 923 SerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu----- 937
Db 373 CTGTGGAGGCGTGGGATTCAGTAAATGACACCGTTCAACCTCAAGCTGATATTATTGAAAG 432
Qy 938 -----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
Db 433 GTTCTCAGTCGGGCGATGATCAACCCAGCGGAGTGGGAGCGCTGAGAGCAACAGC 492
Qy 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
Db 493 GCGGTTCGCCACTTTGAGTATCAGATCGCGGTGAGCTGTGATGACTACTATGCTT 552
Qy 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
Db 553 -----GGCTGTAAAT 561

	:::			:::			2214
D _b	2155	TGCATGTGCTCCTGGCGCTGGGAAGGAACAACCTGTAAACATAGCCCGAACAAGTAGTGTGC					2214
Q _y	1692	ValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyralaGlySerThr					1711
D _b	2215	CTGCCCAACCCTGCCATAATGG-					2238
Q _y	1712	ValThrTyrGlnCysAsnAsnGlyTyTYrLeuLeuGlyAspSerArgMetPheCysThr					1731
D _b	2239	- - - - - GGCACATGTGTGTCTCAACGCGGAGTCCTTTACGTGGCTC					2277
Q _y	1732	AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal					1751
D _b	2278	TGCAGGAGCGCTGGAGGG-					2319
Q _y	1752	GlySerAspCysSerGluHis-					1766
D _b	2320	- - - AATGACTGCAGCCCTCATCCCTGTTACACAGCGGCACCTGTGTGGATGGAGACAAC					2376
Q _y	1767	SerTyrtileCysSerCysValProProTyrThrGlyAsp-					1779
D _b	2377	TGTTACCGGTGCGAATGTGCCCGGGTTTTGCTGGCGCGAGCTGCAGAAATAAACATCAAT					2436
Q _y	1780	- - - - - GlyLysAsnCysAlaGluProIle-					1787
D _b	2437	GAATGCCAGCTTTCACCTTGCTGGCGGACGACCTGTGTGGATGAGATCAATGGCTAC					2496
Q _y	1788	LysCysLysalaProGlylAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal					1807
D _b	2497	CGGTGTCTCGCCCT- - - - - CCAGGGCACAGTGGTGCCAAG-					2532
Q _y	1808	GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle					1827
D _b	2533	- - - - - TGCAGGA-- - - - - GTTTCAGGGAGA					2553
Q _y	1828	ThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro					1840
D _b	2554	CCTTGATCATTCCAGGGAGT- - - - - GTGATACCA					2583

RESULT 13

US-08-400-159-5
Sequence 5, Application US/08400159
Patent No. 5869282
GENERAL INFORMATION:
APPLICANT: Isb-Horowicz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,159
FILING DATE: 07-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie


```
QY 1732 AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
DB 2678 TGCAGGAGGCTGGAGGGG-----CCCATCTGTGTCAGATACC----- 2719
QY 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
DB 2720 ---AATGACTGCAGCCCTCATCCCTGTTACACAGCGCCACCTGTGTGGATGGAGACAAC 2776
QY 1767 SerTrpCysSerCysValProProTrpTrpGlyAsp----- 1779
DB 2777 TGTACCGGTGGATGTGCGCGGTTTGTCTGGGCGGAGCTGCAGATAAACATCAAT 2836
QY 1780 -----GlyAsnAsnCysAlaGluProIle----- 1787
DB 2837 GAATGCCAGCTTCCACCTGTGCTTGGAGCGACCTGTGTGGATGAGATCAATGGCTAC 2896
QY 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTrpVal 1807
DB 2897 CGGTGTGTGCTGCT-----CCAGGGCACAGTGTGTGCCAAG----- 2932
QY 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
DB 2933 -----TGCAGGAA-----GTTTCAGGGAGA 2953
QY 1828 ThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
DB 2954 CCTGTGCATCACCATTGGGGAGT-----GTGATACCA 2983

RESULT 14
US-08-611-729A-5
; Sequence 5, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6464 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 371..4024
US-08-611-729A-5
Alignment Scores:
Pred. No.: 5,27e-57 Length: 6464
Score: 745.00 Matches: 257
Percent Similarity: 34.95% Conservative: 125
Best Local Similarity: 23.51% Mismatches: 349
Query Match: 7.29% Indels: 362
DB: 3 Gaps: 48

US-09-977-053-6 (1-1842) x US-08-611-729A-5 (1-6464)
QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
DB 533 TGCAGCGCGCGCCGGAACCCCGGAGACCGCAAGTGCACCCCGCAGAGTGTGACACATAC 592
QY 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
DB 593 TTCAAGTGTGCTCAAG-----GAGTATCACTCCCGGTCCAGCGCGCGGCCCC 643
QY 869 GlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspPheLeuAspThrVal 888
DB 644 TGCAGCTTCGGCTCAGGG----- 661
QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
DB 662 -----TCCAGCGCTGTATCGGGGGCAACACCTTCAACCTCAAGGCCAGCGCGGC 712
QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
DB 713 AACGACCGCAACCGCATCGTCTTTCAGTTTCGGCTGCCAGGTCCTATACGTTG 772
QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu----- 937
DB 773 CTGTGGAGGGGTGGATTCAGTAATGACACCGTCAACCTGACAGTATTATTGAAAG 832
QY 938 -----AsnGlnArgLeuLeuGlnThrIleGluThrIleThr 950
DB 833 GTTCTCTCACTCGGCGCATCATCAACCCAGCGGAGTGCAGACGCTGAAGCAGAACACG 892
QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
DB 993 GCGGTGCCCCACTTTGAGTATCAGATCCGCTGACCTGTGATGACTACTACTATGCTTT 952
QY 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
DB 953 -----GCTCTAAT 961
QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
DB 962 AAGTTCTGCGCGCCCGCAGAGATGACTTCTTTGGA----- 994
QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysAspGlyIleGlySerTyrGlnAsp 1025
DB 995 -----CACTATCCCTGTGAC-----CAGATGGCCACAAACTTGC 1030
QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
DB 1031 ATGGAAGGCTGATGGGCCCCGGAATGTAACAGAGCTATTTCGCGACAGAGTCTCT 1090
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 1091 AAGCATGGGCTCTTGCAAACTCCAGAGTACTGTCAGGTGTCAG-----TAC 1135
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
DB 1136 GGCTGGCAAGGCGCTG---TACTGTGATAGTGCATCCACCCGGGATGCGTCCACGGC 1192
```

QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
DB 1193 ATCTGTAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGGCGGCAGCTCTGT--- 1249
QY 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
DB 1250 -----GACAAAGATCTCAATTAAGTCTGTGG 1273
QY 1108 Val-----ProCysProGluGly-----LysPheSer 1116
DB 1274 ACTCATCAGCCGTGTCTCAACGGGGGAACTGTAGCAACACAGCCCTGCACAAATATCAG 1333
QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
DB 1334 TGTTC-----TGCCTGTAGGGGTATTTCAGGACCCCACTGT--- 1369
QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSer 1156
DB 1369 ----- 1369
QY 1157 IleThrGluCysSerSerPheSerThrPheSerAlaAlaGluSerValPro 1176
DB 1369 ----- 1369
QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
DB 1370 -----GAAATGTCTGAG-----CACGCC 1387
QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
DB 1388 TGCCTCTCTGATCCCTGTCTACACAGAGCAGCTGTAAAGAGACCTCCCTGGCTTTGAG 1447
QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
DB 1448 TGTGAGTGTCCCAAGCTGGACCGCGCCACATGCTCTACAAACATTGATGACTGTCT 1507
QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
DB 1508 CCTAATAAGTGTCTCCACGGGGCCACTGCCAGGACCTGGTTAAACGGATTAAAGTGTGTG 1567
QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1276
DB 1568 TGCCCCCAGAGTGGGTAAGAACTGCCAGTTAGTAGCAATGAATGTGAGGCCAA 1627
QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
DB 1628 CTTGTGTAAACGCCAAATCCTGTGAAGAATCTCAATTGCCAGCTACTACTGCCAGCTGTCT 1687
QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
DB 1688 CCGGCTGGATGGGTGAGAAATGTGACATAAATATTAATGACTGCTTGGCCAG---TGT 1744
QY 1317 LeuAsnAsnAlaValCysGluAspGlnValGlyPheLeuCysLysCysProProGly 1336
DB 1745 CAGATGACGCTCTCTGTGGGATTTGGTTAATGTTATGCTGTATCTGTCTCACTGCG 1804
QY 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
DB 1805 TATGCAGCGGATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAAT 1864
QY 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
DB 1865 GGGGTGCACTGTGAGAAATCAACAGATTCAGTGTGTGTGCCACTGGTTTCTCT 1924
QY 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
DB 1925 GGAACCTCTGTCTACATGACATGATTAATTTGTAGCCTTAATCCTTGCAGAACCGTGC 1984
QY 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
DB 1985 CAGTGTCTACACCGTGGCAGTGACTATTCTTGTCAAGTGGCCCCGAGGACTATGAGGGCAAG 2044

QY 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
DB 2045 AACTGCTCACCTGAAAGACCCACTGCCGCGACAGCCCTCTGTGAAGTG----- 2092
QY 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
DB 2093 -----ATTGAC-----AGCTGCACAGTGGCC 2113
QY 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
DB 2114 ATGGCTTCCAAGAC-----ACACCTGAAGGGTGGGTATATTTCTCTCC 2158
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
DB 2159 AAC-----GTCTGTGTCTCTCAC 2176
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAla 1514
DB 2177 GGGAG-----TGCAAGAGTCACTGGGAGGCAATTC----- 2209
QY 1515 IleThrTyrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
DB 2210 ---ACCTGTGACTGTAAACAAAGGCTTCACGGGAACATACTGCCATGAAATATTAAATGAC 2266
QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyAlaLeuValLeuGly 1554
DB 2267 TGTGAGAGCAAC-----CCTGTAGAAACGGTGGCACTTGCATCGATGCT 2311
QY 1555 GlnGluGlnAspLysGlyGlyGlyPheSerProAlaGluSerPheValGlySerIle 1574
DB 2312 GTCAACTCTTACAAG-----TGCACTGTAGTAGCGCTGGGAGGGGCTTAC 2359
QY 1575 SerGlnLeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
DB 2360 TGTGAACCAATATTATTAATGACTGCCAGCCAGAACCC----- 2395
QY 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
DB 2396 ---TGCCACAT----- 2404
QY 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
DB 2405 -----GGGGGCACGCTGCCGACCTGTGTCAATGACTTCTACTGTGACTGTAAATATGGG 2458
QY 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
DB 2459 TGAAGGAAGAACCTGCCACTCAGTGACATGAGTGTGTGAGGCCACGTCGCAACAC 2518
QY 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
DB 2519 GGTGGCACCTGTATGATGAGGGGGATGCTTTTAAG----- 2554
QY 1672 CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
DB 2555 TGCATGTGCTCTGGCGGCTGGGAAGAACCACTGTAAACATAGCCCGAACAACAGTAGCTGC 2614
QY 1692 ValProProProLeuAsnGlyPheHisSerAlaAspPheTyrAlaGlySerThr 1711
DB 2615 CTGCCCAACCCCTGCCATAATGG----- 2638
QY 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
DB 2639 -----GGCAGATGTGTGGTTCACGGCGAGTCTCTTACGTGGTC 2677
QY 1732 AspAsnGlySerTyrPheAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
DB 2678 TGCAAGGAAGGCTGGGAGGG-----CCCATCTGTGTCTCAGAAATACC----- 2719
QY 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
DB 2720 ---AATGACTGAGCCCTCATCCCTGTTTCAACAGCGGCACCTGTGTGTGATGAGACAC 2776
QY 1767 SerTyrIleCysSerCysValProTyrThrGlyAsp----- 1779

```
Db 2777 TGTACCGTGCAGATGTGCCCGGGTTTCTGTCGGCCGACTGCAGAAATAACATCAAT 2836
Qy 1780 -----GlyAsnAspCysAlaGluProIle----- 1787
Db 2837 GAATGCCAGCTTCCACCTGTGCTTGTGAGCGACCTGTGTGGATGAGATCAATGGCTAC 2896
Qy 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleThrVal 1807
Db 2897 CGGTGTGTCTCCCT-----CCAGGGACAGTGTGTGCCAAG----- 2932
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
Db 2933 -----TGCCAGGA-----GTTTCAGGGAGA 2953
Qy 1828 ThrCysLeuGluSerGlyGluThrAsnHisLeuIlePro 1840
Db 2954 CCTTGCATCACCATGGGAGT-----GTGATACCA 2983

RESULT 15
US-09-068-740A-10
; Sequence 10, Application US/09068740A
; Patent No. 6337387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; US-09-068-740A-10

Alignment Scores:
Pred. No.: 3,01e-57 Length: 4208
Score: 744.00 Matches: 257
Percent Similarity: 34.95% Conservative: 125
Best Local Similarity: 23.51% Mismatches: 349
Query Match: 7.28% Indels: 362
DB: 4 Gaps: 48

US-09-977-053-6 (1-1842) x US-09-068-740A-10 (1-4208)
Qy 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
Db 571 TGCAGCGCGCGCGGAGACCCCGGGAGACCGCAAGTGCACCCCGCAGAGTGTGACACATAC 630
Qy 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
Db 531 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGGTCCAGCGCGCGGGGCCCC 681
Qy 869 GlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspPheLeuAspThrVal 888
Db 682 TGCAGCTTCGCTCAGG----- 699
Qy 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
Db 700 -----TCCAGCGCTGTGATCGGGGGCAACACCTTCAACAGGCCAGCGCGGCG 750
```

```
Qy 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
Db 751 AACGACCGCAACCGCATGTGCTGCTTTTCAGTTTTCGCTGCGCGAGGTCCTATACGTTG 810
Qy 923 SerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu----- 937
Db 811 CTGTGGAGGGGTGGGATTCAGATTAATGACACCGTTCAACCTGCACAGTATTATTGAAAG 870
Qy 938 -----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
Db 871 GCTTCTCCTCTGGGCAATCAACCCAGCGCGAGTGGCAGCGCTGAAGCAGAACAG 930
Qy 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
Db 931 GCGGTGCGCCCTTTGAGTATCAGATCCGCGTGACCTGTGATGACTACTACTATGCTTT 990
Qy 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
Db 991 -----GGTGTCAAT 999
Qy 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
Db 1000 AAGTTCTGCGCGCCAGAGATGACTTCTTTGGA----- 1032
Qy 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
Db 1033 -----CACTATGCTGTGAC-----CAGAATGGCAACAAACTTGC 1068
Qy 1026 GluGluGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
Db 1069 ATGGAAGGCTGGATGGCGCCGCAATGTAAACAGAGCTATTTCGCAAGGTCGACGTCCT 1128
Qy 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
Db 1129 AAGCATGGTCTTGCAAATCCAGGTGCTGACAGTGCAG-----TAC 1173
Qy 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
Db 1174 GGCTGGCAAGGCGCTG---TACTGTGATAAGTGCATCCACACCCGCGGATGCGTCCAGCGC 1230
Qy 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
Db 1231 ATCTGTAAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGGCGGCGCTGTGT--- 1287
Qy 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
Db 1288 -----GACAAAGATCTCAATTACTGTGGG 1311
Qy 1108 Val-----ProCysProGluGly-----LysPheSer 1116
Db 1312 ACTCATCAGCGCTGTCTCAACGGGGGAACCTTGTAGCAACACAGCGCCCTGACAAATATCAG 1371
Qy 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
Db 1372 TGTTC-----TGCCTGAGGGGTATTTCAGGACCCCAACTGT--- 1407
Qy 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
Db 1407 ----- 1407
Qy 1157 IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluSerValValPro 1176
Db 1407 ----- 1407
Qy 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
Db 1408 -----GAAATGCTGAG-----CACGCC 1425
Qy 1197 CysPheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
Db 1426 TGCCTCTCTGATCCTCTCACACAGAGCGAGCTGTAGGAGACCTCCCTCGGCGCTTGGAG 1485
Qy 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
```

Db 1486 TGTGAGTGTTCCTCCAGGCTGGACCGGCCCATGCTCTACAAACATTGATGACTGTCT 1545
Qy 1237 ProLeuProCysLeuAsnAsnGlyValCysAspLeuValGlyGluPheIleCysGlu 1256
Db 1546 CCTAATACTGTCCACGGGGGACCTGCCAGGACCTGTGTTAACGGATTAAGTGTGTG 1605
Qy 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1276
Db 1606 TGGCCCCACACGTGACCTGGGAACAGTGCAGTTAGATGCAATGAATGTGAGGCGCAA 1665
Qy 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
Db 1666 CTTGTGTAAACGCGCAATCTCTGTAAGAATCTATTGCCAGCTACTCTGCGACTGTCTT 1725
Qy 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
Db 1726 CCGGCTGGATGGGTGAGATTTGTCATTAATTAATGACTGCTTGGCCAG--TGT 1782
Qy 1317 LeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysGlyCysProGly 1336
Db 1783 CAGATGACGCTCTGCTGGGATTTGGTTAATGTTATCGTGTATCTGTCTCCACTGCG 1842
Qy 1337 PheLeuGlyThrArgCysGlyAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
Db 1843 TATGACGGCGATCACTGTGAGAGACATCGATGAATGTCGACCAACCCCTGTTTGAAT 1902
Qy 1357 GlyAlaThrCysLysGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
Db 1903 GGGGTCACTGTGAGATGAATCAACAGATTCAGTGTCTGTGTCCACTGTTCTCT 1962
Qy 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
Db 1963 GCAAACTCTGTGACGTGACATCGATTAATTTGTGAGCCTAATCCCTGCCAGAACGGTGC 2022
Qy 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGln 1416
Db 2023 CAGTCTCAACCGTGCAGTGAATCTTCTGCAAGTGCCTGAGGACTATGAGGCGCAG 2082
Qy 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGlyValSerGlyIleTyr 1436
Db 2083 AACTGCTCACCTGAAGACCACTGCCCGCAGCCCTCTGTGAGTGTG-- 2130
Qy 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThr 1456
Db 2131 -----ATTGAC-----AGCTGCACAGTGGCC 2151
Qy 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
Db 2152 ATGGCTTCCACGAC-----ACACCTGAAGGGGTGCGGTATATTTCCTCC 2196
Qy 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
Db 2197 AAC-----GTCTGTGCTCTCTAC 2214
Qy 1495 GlyArgGlyLysIleThrAsnCysProSerValAsnAspGlyArgThrPheHisIleAla 1514
Db 2215 GCGAAG-----TGCAAGAGTCACTGGGAGGCAAAATTC----- 2247
Qy 1515 IleThrThrThrSerAlaAsnGlyIleThrLysValTyrIleAspGlyLysLeuSerAsp 1534
Db 2248 ---ACCTGTGACGTGAACAAGGCTTCACGGGAACATATGCGCATGAAATAATTATGAC 2304
Qy 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyAlaLeuValLeuGly 1554
Db 2305 TGTGAGAGCAAC-----CCTGTAGAAACGGTGGCACTTGCATCGATGT 2349
Qy 1555 GlnGluGlnAspLysGlyGlyPhePheSerProAlaGluSerPheValGlySerIle 1574
Db 2350 GTCAACTCTCTACAG-----TGCAATCTGTAGTACGCGGTGGGAGGCGGCTAC 2397
Qy 1575 SerGlnLeuAsnThrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
Db -----

Db 2398 TGTGAACCAATATTATTGACTGACGAGCCAGAACCCC----- 2433
Qy 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
Db 2434 ---TGCCACAAT----- 2442
Qy 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
Db 2443 -----GGGGGACACCTGCGGACCTGTGTCATGACTTCTACTGTGACTGTAAAAATGGG 2496
Qy 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
Db 2497 TGAAGAAGGAAGACCTGCGACCTCACTGACACTGATGTGTGAGCCACGTCGCAACAC 2556
Qy 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
Db 2557 GGTGGCACCTGTATGATGAGGGGATGCTTTTAAG----- 2592
Qy 1672 CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
Db 2593 TGCATGTGTCTGCGCGCTGGGAAGAAACACTGTAAACATAGCCCGAAACAGTAGCTGC 2652
Qy 1692 ValProProProLeuGluAsnGlyPheHisSerAlaAspPheTyrAlaGlySerThr 1711
Db 2653 CTGCCCAACCCCTGCCATAATGG----- 2676
Qy 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
Db 2677 -----GGCACATGTGTGTGTCACACGGCGAGTCTTTTACGTGCGTC 2715
Qy 1732 AspAsnGlySerTyrAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
Db 2716 TGCAAGGAAGGCTGGAGGGG-----CCCATCTGTGTCTGCAATATACC----- 2757
Qy 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
Db 2758 ---ATGACTGACGCGCTCATCCCTGTACACACGGCGACCTGTGTGATGGAGACAC 2814
Qy 1767 SerTyrIleCysSerCysValProTyrThrGlyAsp----- 1779
Db 2815 TGGTACCGTGGCAATGTCCCGCGGTTTGTGCGCGGCTGCGAGATAAATCAAT 2874
Qy 1780 -----GlyLysAsnCysAlaGluProIle----- 1787
Db 2875 GAATCCAGTCTTACCTTGTGCGGACCTGTGTGATGAGATCAATGGCTAC 2934
Qy 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
Db 2935 CGGTGTGTGCGCT-----CCAGGGCACAGTGTGTCACAG----- 2970
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
Db 2971 -----TGCCAGGAA-----GTTTCAGGAGAGA 2991
Qy 1828 ThrCysLeuGluSerGlyGluTyrAsnHisLeuLeuPro 1840
Db 2992 CCTTGATCATTGCGGAGT-----GTGATACCA 3021
RESULT 16
; US-08-882-046-1
; Sequence 1, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Kwantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 FILING DATE: 25-JUN-1997
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-UW 2637
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5590 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 414..4068
 US-08-882-046-1

Alignment Scores: 5,01e-57 Length: 5590
 Pred. No.: 744.00 Matches: 257
 Score: 34.95% Conservative: 125
 Best Local Similarity: 23.51% Mismatches: 349
 Query Match: 7.28% Indels: 362
 DB: 3 Gaps: 48

US-09-977-053-6 (1-1842) x US-08-882-046-1 (1-5590)

QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
 DB 576 TGGCGGGCCCGGAACCCGGGAGCCGAGTGACACCGCGAGGTGTGACACATAC 635
 QY 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
 DB 636 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGCGTCAACGCGCGGGGCGCC 686
 QY 869 GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspPheLeuAspThrVal 888
 DB 687 TGCAGCTTCGGCTCAGGG----- 704
 QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
 DB 705 -----TCCAGCGCTGTGTCATCGGGGGCAACACCTTCACTCAAGCGCGCGCGG 755
 QY 909 LeuSerAspTyrIleLysLeuIlePheAsnIle-----ThrAla 922
 DB 756 AACGACCGCAACCGCATCGTGTGCTTTCAGTTTCGCTTCGCGCGAGGTCTATACGTTG 815
 QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluThrGlu----- 937
 DB 816 CTTGTGAGCGGTGGGATTCCAGTAATGACACCGTTCACTGACAGTATTATTGAAAG 875
 QY 938 -----AsnGlnIleArgLeuLeuGlnThrLeuGluThrIleThr 950
 DB 876 GCTTCTCACTCGGGCATGATCAACCCAGCGGAGTGGCAGACGCTGAGACGAGAACG 935
 QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
 DB 936 GGGGTTGCCCATTTTGAGTATCAGATCCGCGTGACCTGTGATGACTACTACTATGGCTTT 995

QY 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
 DB 996 -----GGCTGCAAT 1004
 QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
 DB 1005 AAGTTCTGCGGCCCCAGAGATGACTTCTTTGGA----- 1037
 QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
 DB 1038 -----CACTATGCTCTGAC-----CAGATGGCAACAAACTTGC 1073
 QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
 DB 1074 ATGGAGGCTGGATGGGCCCCGATGTAACAGAGCTATTTCGCGACAGGCTGCAGTCT 1133
 QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
 DB 1134 AAGCATGGGTCTTGCAAACTCCAGGTGACTGCAGGTGCCAG-----TAT 1178
 QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
 DB 1179 GGCTGGCAAGGCGCTG---TACTGTGATAAGTGCATCCACACCGGGGATGGCTCCACGC 1235
 QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
 DB 1236 ATCTGTAATGAGCCCTGGCAGTGCCTCTGTGAGACCACTGGGGCGGCGCTGT--- 1292
 QY 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
 DB 1293 -----GACAAAGATCTCAATTACTGTGG 1316
 QY 1108 Val-----ProCysProGluGly-----LysPheSer 1116
 DB 1317 ACTCATCAGCGGTGTCTCAACGGGGGAACCTGTAGCAACACAGGCCCTGACAAATATCAG 1376
 QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
 DB 1377 TGTTC-----TGCTCTGAGGGGTATTTCAGGACCCCACTGT--- 1412
 QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
 DB 1412 ----- 1412
 QY 1157 IleThrCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValPro 1176
 DB 1412 ----- 1412
 QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
 DB 1413 -----GAAATTGCTGAG-----CACGCC 1430
 QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
 DB 1431 TGCCTCTCTGATCTCTGTCACACAGGCGAGCTGTAAAGGAGACCTCCCTGGGCTTTGAG 1490
 QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
 DB 1491 TGTGAGTGTTCCTCCAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTCT 1550
 QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
 DB 1551 CCTAATAACTGTTCCACAGGGGGCACCTGCCAGGACCTGGTTAAGGATTTAAGTGTGTG 1610
 QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer 1276
 DB 1611 TGGCCCCCAGTGGAGTGGGAAAACGTCGCGAGTTAGTCAAAATGATGTGAGGCCAAA 1670
 QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
 DB 1671 CCTGTGTAAACGCCAAATCCTGTAAAGAAATCTATTGCCAGCTACTACTGACACTGTCTT 1730

Qy 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
Db 1731 CCCGGCTGGATGGTCACAAATTTGTCATAAATAATTAATGACTGCTTGGCCAG---TGT 1787
Qy 1317 LeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysValCysProProGly 1336
Db 1788 CAGAAATGACCGCTCTCTGGGGATTTGGTTAATGTTATCCCTGTATCTGTCCACTGCTG 1847
Qy 1337 PheLeuThrArgCysGlyLeuAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
Db 1848 TATCGACGGCATCATCTGTGAGAGACATCATGATGATGATGATGATGATGATGATGAT 1907
Qy 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
Db 1908 GGGGGTCACTCTGCAATGAATAACACAGATTCCAGTGTCTGTGTCCCACTGGTTCTCT 1967
Qy 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
Db 1968 GGAACCTCTGTGCTGACGTGACATCATGATTTATGTCAGTCTATCTCTGACAGACGCTCC 2027
Qy 1397 ThrCysValAspGluLeuAsnSerTySerCysLysCysGlnProGlyPheSerGlyGln 1416
Db 2028 CAGTGTACAAACCGTCCAGTCACTATTCTGCAAGTSCCCGAGGACTATGAGGCAAG 2087
Qy 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTy 1436
Db 2088 AACTGCTCACACCTGCAAGACCACTGCGCCGACGACCCCTCTGTGAAGTG----- 2135
Qy 1437 GlyTyValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrp 1456
Db 2136 -----ATTGAC-----AGCTGCACAGTGCC 2156
Qy 1457 MetLysSerAspAspMetAsnTyGlyThrPro-----IleSerTyAlaValAsp 1474
Db 2157 ATGGCTTCAACGAC-----ACACCTGAAGGGGTGCGGTATATTTCCTCC 2201
Qy 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyAsnGlyTyrValLeuTyValAsn 1494
Db 2202 AAC-----GTCTGTGTCTCTCAC 2219
Qy 1495 GlyArgGluTyIleThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAla 1514
Db 2220 GGGGAG-----TGCAAGAGTCACTCGGAGGCAATTC----- 2252
Qy 1515 IleThrTrpThrSerAlaAsnGlyIleThrLysValTyIleAspGlyLysLeuSerAsp 1534
Db 2253 ---ACCTGTGCTGTGACAAAGGCTTCACGGGACATATGCGCATGAAATAATTAATGAC 2309
Qy 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGly 1554
Db 2310 TGTGAGAGCAAC-----CCTTTGAGAAACGGTGGGCACTTGCATCATGATGCT 2354
Qy 1555 GlnGluGlnAspLysGlyGlyGlyPheSerProAlaGluSerPheValGlySerIle 1574
Db 2355 GTCAACTCCTACAG-----TGCACTCTGTAGTACGGCTGGAGGGGCGCTTAC 2402
Qy 1575 SerGlnLeuAsnLeuTrpAspTyValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
Db 2403 TGTGAACCAATATTAATGACTGACGACGACGACGACGACGACGACGACGACGACGAC 2438
Qy 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGly 1614
Db 2439 ---TGCCACAAAT----- 2447
Qy 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
Db 2448 -----GGGGGACCTGTGCGGCACTTGTCAATGACTTCTACTGTGACTGTAAATAATGGG 2501
Qy 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
Db 2502 TGGAAAGGAAGACCTGCCACTCAGTGCATGATGATGATGATGATGATGATGATGATGAT 2561
Qy 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTy 1671

Db 2562 GGTGGCACCTGCTATGATGAGGGGATGCTTTTAAG----- 2597
Qy 1672 CysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
Db 2598 TGCATGTGTCTCTGGCGCTGGGAGGAACAACCTGTAAACATAGCCCGAAACAGTAGCTGC 2657
Qy 1692 ValProProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyAlaGlySerThr 1711
Db 2658 CTGCCCAACCCCTGCATAATGGG----- 2681
Qy 1712 ValThrTyGlnCysAsnAsnGlyTyTyTyLeuLeuGlyAspSerArgMetPheCysThr 1731
Db 2682 -----GGCACATGTGTGTGTCAACGGGAGTCTTTTACGTGGCTC 2720
Qy 1732 AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
Db 2721 TGCAAGGAAGGCTGGAGGGG-----CCCATCTGTGTCTCAGATATACC----- 2762
Qy 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
Db 2763 ---AATGACTGCGCCCTCATCTCCCTTTACACACGGGCACTGTGTGTGATGGAGACAAC 2819
Qy 1767 SerTyIleCysSerCysValProProTyTyThrGlyAsp----- 1779
Db 2820 TGTACCGGTGGATGTGCTCCCGGTTTGTGTCGCGGCTGCAGATATAACATCAAT 2879
Qy 1780 -----GlyLysAsnCysAlaGluProIle----- 1787
Db 2880 GAATGCCAGTCTTACCTTGTGCTTGTGAGCGACTGTGTGATGATGATCAATGGCTAC 2939
Qy 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyThrVal 1807
Db 2940 CGGTGTGTCTGCTCCCT-----CCAGGGCACAGTGTGTGCCCAAG----- 2975
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyGlnLeuMetGlyValThrLysIle 1827
Db 2976 -----TGCCAGGAA-----GTTTCAGGGAGA 2996
Qy 1828 ThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 2997 CCTTGATCACCATTGGGGAGT-----GTGATACCA 3026

RESULT 17

US-08-400-159-9

; Sequence 9, Application US/08400159

; Patent No. 5869282

; GENERAL INFORMATION:

; APPLICANT: Ish-Horowitz, David

; APPLICANT: Henrique, Domingos M.P.

; APPLICANT: Lewis, Julian H.

; APPLICANT: Myat, Anna M.

; APPLICANT: Fleming, Robert J.

; APPLICANT: Artavanis-Tsakonas, Spyridon

; APPLICANT: Mann, Robert S.

; APPLICANT: Gray, Grace E.

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/400,159

FILING DATE: 07-MAR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Miarock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 7326-029
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3582 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..3582
 US-08-400-159-9

Alignment Scores:
 Pred. No.: 2,03e-52 Length: 3582
 Score: 690.00 Matches: 238
 Percent Similarity: 36.34% Conservative: 113
 Best Local Similarity: 24.64% Mismatches: 301
 Query Match: 6.75% Indels: 315
 DB: 2 Gaps: 48

US-09-977-053-6 (1-1842) x US-08-400-159-9 (1-3582)

QY 987 PheCysArgProGlySerValLeuArgGlyMetCysValAenCysProLeuGlyThr 1006
 DB 517 TTTTGTGCGCCGAGAGT----- 534
 QY 1007 TyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGlu 1026
 DB 535 -----GACTTCTTCACTCACCATACCTGT-----GACCAG 564
 QY 1027 GluGlyGlnLeuGluCys-----LysLeuCys 1035
 DB 565 AATGGCAACAAACCTCTTGGAGGCTGGACCGGACCAAGATGCACAAAGCTATTGT 624
 QY 1036 ProSerGlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGln 1055
 DB 625 CGTCAGGAGTGTAGCCCAACAGATGGTCTTGCACAGTTCAGGAGAGTGCAGTGTCTAG 684
 QY 1056 CysLysGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCys----- 1072
 DB 685 -----TATGGATGCAAGGC---CAGTACTGTGATAGTGCATTCACAC 726
 QY 1073 ProLeuGlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsn 1092
 DB 727 CCGGGATGTGTCATGGCACTTGCAITGGAACCATGCGCAGTGCCTC---TGTGAACCAAC 783
 QY 1093 ThrSerThrValLysArgGlyAlaVal-----AsnIleSerAlaCysGlyVal--- 1108
 DB 784 TGGGGT-----GGTCAGCTCTGTGACAAAGACCTGAACTACTGTGGAAACCCAC 831
 QY 1109 ---ProCysProGluGly-----LysPheSerArgSer 1118
 DB 832 CCACCTCTTTGAATGTTGGTGTACCTGCAGCAACACTGCCGCCGATATAATACCAGTGTCC 891
 QY 1119 GlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAla-----Gly 1136
 DB 892 -----TGCCTCTGAGGGTTACTCAGGACAGAACTGTGAAATAGCG 930
 QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
 DB 931 GAGCATCGGCTCTCTGATCCGTGCCACAC-----GAGGAGGC----- 972
 QY 1157 IleThrGluCysSerPheSerSerThrPheSerAlaAlaGluSerValPro 1176

DB 973 -----TGCCTAGAAAACGCTCTACAGATTT-----GAATGTGTGTGTCGA 1011
 QY 1177 ProAlaSerLeuGlyHisIleLysArgHisGluIleSerSerGlnValPheHisGlu 1196
 DB 1012 CTTGGCTGGGCTGGA-----CCAATTCGCACTGATAATATTGATGAT 1053
 QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
 DB 1054 TGTTCCTCAAAATCCCTGTGTGTCATGGAGGAACCTGCCCAAGATCTAGTTGATGGATTAAAG 1113
 QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
 DB 1114 TGTATTTCCTCCCTCCTCAGTGGACTGGCAAAACATGCCAGCTAGATGCGAATGAATGTGAG 1173
 QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
 DB 1174 GGCAAACCTCTGTGTCAATGCCACTCTCGAGGAACCTTGTATGGAGCTACTATTGTGAC 1233
 QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer 1276
 DB 1234 TGCATTACTGCTGTGCTGGCCCAACTGTGTGATATAATAATTATTGATTGTGCTGGACAA 1293
 QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
 DB 1294 ---TGTGAGATGGAGGATCCTGTGGGACTTGGTTAATGGTTATCGTGTGATCTGTTCA 1350
 QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
 DB 1351 CTTGGCTATGCGAGGAGATCACTGTGAGAAAGACATCAATGAATGTGCAAGTAACCTTGC 1410
 QY 1317 LeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProGly 1336
 DB 1411 ATGAATGGGGTCACTGCCAGATGAATCAATGGATTCCAATGTCTGTCTCTCTGCT 1470
 QY 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
 DB 1471 TTCTCAGGAACCTCTGTGCTGCTGATATAGACTACTGTGAGCCAAACCTTGCAGAAC 1530
 QY 1357 GlyAlaThrCys-----LysAsp----- 1360
 DB 1531 GGTGCCAGCTCTTCAATCTTGTATGCACTATTCTGTAACTGCCCTGAAGATTACGAA 1590
 QY 1361 -----LysAsp----- 1362
 DB 1591 GGCAAGAAAGTCTCCACCTCGAAAGATCACTGCGCACAACTCTTGTGAAAGTAAATCGAC 1650
 QY 1363 -----GlyAlaAsnSer----- 1366
 DB 1651 AGCTGTACAGTGGCAGTGGCTTCTTAACAGCACACACAGAGGAGTTCGTTCATTTCTTCA 1710
 QY 1367 -----PheArgCysLeu 1370
 DB 1711 AATGTCTGTGTCTCTCATGGAAATGCAGAGCCAGCGGTGGAAATTCACCTGTGNA 1770
 QY 1371 CysAlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsn 1390
 DB 1771 TGCAACAAAGGATTCTCTGCGCACCTACTGTCAATGAGAATAATCAATGATGCTGAGAGCAAC 1830
 QY 1391 ProCysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGln 1410
 DB 1831 CCCTGTAAAAATGGTGGCAGCTTGTATTGACCGGTGTAACTCTTCAAAATGTTATTGTAGT 1890
 QY 1411 ProGlyPheSerGlyGlnArgCysGluThrGluGlnSer-ThrGlyPheAsnLeuAspPh 1430
 DB 1891 GATGATGGGAGGAGACATATTGTGAACAATAATTATGACTGCGAGTAAACCCCT--- 1948
 QY 1430 eGluValSerGlyIleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAl 1450
 DB 1949 -----GCCACAATGGAGGA 1962
 QY 1450 aLeuThrCysThrPheTrpMetLysSerSerAspAspMetAsnTyrGlyThrProIleSe 1470

Db 1963 ACTGCCAGACTTGGTCAATGACTTCTCTCTGTAATGTAATAATGGGT----- 2011
Qy 1470 rTyralaValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrVa 1490
Db 2011 ----- 2011
Qy 1490 lLeuTyrValAsnGlyArgGluValleThrAsnGlyProSerValAsnAspGlyArgTr 1510
Db 2012 -----GGAAGGAAACCTTGCCTCTCTGACGACGCGAGTGTGATGAGGA 2058
Qy 1510 phis-----HisleAlaileThrTrpThrSerAlaAsnGlyIleTrpLy 1525
Db 2059 ACATGCAATAATGGAGGACATGTTATGATGAGGGGACACTTTCAGTGCATGT----- 2113
Qy 1525 sValTyrleAspGlyLys-----LeuSerAspGlyGlyAlaGlyLeuSerValGl 1542
Db 2114 -GTCTGCAGGATGGAAAGAGCCACTTGTATATAGCAAGCAACAGCAGCTGCCTGCCA 2172
Qy 1542 yLeuProIleProGlyGlyValaleuValleuGlyGlnGluGlnAspLysLysGlyGl 1562
Db 2173 AACCCCTGTAC-AATGGTGTGACTGTGTAGTT-----AGTGGGA 2213
Qy 1562 uGlyPheSerProAla-----GluSerPheValGlySerIleSerGlnLeuAsnLeuTr 1580
Db 2214 TCTTTTCACTGTGTCGACGAGGGCTGGAGGACCGACATGACTACTCAGACACAA 2273
Qy 1580 pAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGlu----- 1598
Db 2274 TGAC-----TGCAGTCTCATCTTGTATACACAGTGTACT---TGTGTGGATGAGA 2324
Qy 1599 -----GluLeuSerLysGlyAsnValleuAlaTrpProAspPheLeuSe 1613
Db 2325 CAACGTGTACCGCTGTGAGTGGCTCCCGC-----TTCGAGGTCCCGACTGT----- 2373
Qy 1613 rGlyIleValGlyLysValLysValleAspSerLysSerIlePheCysSerAspCysProAr 1633
Db 2374 -----AGCATCAATCAATGAA-----TGTCACTTCCACCTG 2408
Qy 1633 gLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerly 1653
Db 2409 TGCCTTTGGGCT-----ACTGTGTGGATGAAAT-----AATGG 2444
Qy 1653 sValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLe 1673
Db 2445 GTACCGTGTCAATTTGTCACCGGGT----- 2469
Qy 1673 uAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValPr 1693
Db 2470 -----CGCAGTGTCCAGGATCCAGGAAGTTACA---GGCAGGCC 2507
Qy 1693 o-----ProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAl 1708
Db 2508 TTGCTTTACAGTATTCAGAGTAATCCAGACGGTCTAAGTGGGATGATGAC----- 2559
Qy 1708 aGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMe 1728
Db 2560 -----TGTAATACTTGTCACTGTTGATGGA-----AAGT 2591
Qy 1728 tPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspVal----- 1746
Db 2592 CACCTGTTCTAAG-----GTTTGTGTGTCTCTCGACCTGTATATAATACATGCCAAGG 2645
Qy 1747 -----AspGluCysAlaValGlySerAspCys-----SerGluHisAlaSerCy 1761
Db 2646 TCATAATGAATGCCAGCTGGACACGCTGTGTCTGTAAAGAGAACCAT-----TG 2699
Qy 1761 sLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLy 1781
Db 2700 TTTCACTCATCTTGTGCTGAGTGGTGAATGCTGGCCT----- 2739
Qy 1781 sAsnCysAlaGlaProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSe 1801
Db 2740 -TCTAATCAGCAGCCTGTGAGACCAATGCG-----AATTC 2774

Qy 1801 rGlyGluIleTyrThrValGlyAlaAla---ValThrPheSerCysGlnGluGlyTyrGl 1820
Db 2775 TGATTTCTTATTACCAGATAATGTGCCCAATCATCCTTCACTTTAATAGGAATGAT 2834
Qy 1820 nLeuMetGlyValThr 1825
Db 2835 GGCACCAGGCTTACC 2850
RESULT 18
US-08-611-729A-9
; Sequence 9, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Taakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace S.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3579
US-08-611-729A-9
Alignment Scores:
Pred. No.: 2,03e-52 Length: 3582
Score: 600.00 Matches: 238
Percent Similarity: 36.34% Conservative: 113
Best Local Similarity: 24.64% Mismatches: 301
Query Match: 6.75% Indels: 315
DB: 3 Gaps: 48
US-09-977-053-6 (1-1842) x US-08-611-729A-9 (1-3582)
Qy 987 PheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThr 1006
|||

QY 1613 rGlylleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProAr 1633
 Db 2374 -----AGGATCAACATCAATGAA-----TGTCAGTCTTACCCCTG 2408
 QY 1633 gLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLy 1653
 Db 2409 TGCCTTTGGGCT-----ACTGTGTGGATGAAT-----AATGG 2444
 QY 1653 sValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLe 1673
 Db 2445 GTACCGTTGCATTTGTCCACCGGCT----- 2469
 QY 1673 uAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGlyValPr 1693
 Db 2470 -----CGCAGTGTCCAGATGCCAGGAAGTTTACA-----GGAGGCC 2507
 QY 1693 o-----ProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAl 1708
 Db 2508 TTGCTTTTACCAGTATTCAGTAATGCCAGACGGTCTAAGTGGGATGATGAC----- 2559
 QY 1708 aGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrIleuLeuGlyAspSerArgMe 1728
 Db 2560 -----TGTAATACTTCTCAGTGTTCGAATGGA-----AAAGT 2591
 QY 1728 tPheCysThrAspAsnGlySerTyrAsnGlyValSerProSerCysLeuAspVal----- 1746
 Db 2592 CACCTGTCTTAAG-----GTTTGTGTGGTCTCGACCTTGTATAATACATGCCAAAGG 2645
 QY 1747 ---AspGluCysAlaValGlySerAspCys-----SerGluHisAlaSerCy 1761
 Db 2646 TCATAATGAATGCCAGCTGGACAGCTGTGTTCTCTGTTAAAGAAGACCAT-----TG 2699
 QY 1761 sLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLy 1781
 Db 2700 TTTCACTCATCTTGCTGTGCTGAGTGGGTGAATGCTGGCT----- 2739
 QY 1781 sAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSe 1801
 Db 2740 -TCTAATCAGCAGCGCTGTGAAGACCAATGC-----AATTC 2774
 QY 1801 rGlyGluIleTyrThrValGlyAlaAla---ValThrPheSerCysGlnGluGlyTyrGl 1820
 Db 2775 TGATTCATTATACCAGATAATTGTGCCACATCACCTTCACCTTTAATAAGGAATGAT 2834
 QY 1820 nLeuMetGlyValThr 1825
 Db 2835 GGCACAGGCGCTTACC 2850

RESULT 19

US-09-214-278-4
 ; Sequence 4, Application US/09214278
 ; Patent No. 6291210
 ; GENERAL INFORMATION:
 ; APPLICANT: Sakano, Seiji
 ; APPLICANT: Itob, Akira
 ; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
 ; FILE REFERENCE: KP-8576
 ; CURRENT APPLICATION NUMBER: US/09/214,278
 ; CURRENT FILING DATE: 1999-01-26
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 3955
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (12)...(3725)
 ; NAME/KEY: sig_peptide
 ; LOCATION: (12)...(89)
 ; NAME/KEY: mat_peptide
 ; LOCATION: (90)...(3725)
 US-09-214-278-4

Alignment Scores: 7,128-51 Length: 3955
 Pred. No.: 674.00 Matches: 232
 Score: 34.42% Conservative: 86
 Best Local Similarity: 25.11% Mismatches: 276
 Query Match: 6.60% Indels: 330
 DB: 3 Gaps: 45
 US-09-977-053-6 (1-1842) x US-09-214-278-4 (1-3955)
 QY 967 LeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986
 Db 579 CTGGAGCTGCAGATCGGCTGCGTCGACGAGAACTACTACAGCGCCACTTGCACAAG 638
 QY 987 PheCysArgProGlySerValLeuArgIleArgMetCysValAsnCysProLeuGlyThr 1006
 Db 639 TTCTGCGCGCCGCGCAACGACTTTTTCGCG----- 668
 QY 1007 TyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGlu 1026
 Db 669 -----CACTACACCTGCG-----GACCAG 686
 QY 1027 GluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSer 1046
 Db 687 TACGGCAAC-----AAGCCTGCATGGAGCGCTGGATGGGCAAG----- 725
 QY 1047 ArgAsnIleSerAspCysLys---AlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065
 Db 726 -----GAGTGCAGGAAGCTGTGTAAACAAGGCTGTAAATTTGCTCCACGG 773
 QY 1066 LeuGluThrCysGluSerCysProLeu---GlyThrTyrGlnProLysPheGly----- 1082
 Db 774 -----GGATGCACCTGCTGCGGAGTGCAGTGCAGCTACGGCTGGCA 818
 QY 1083 SerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsn 1102
 Db 819 GGGAGTTCTGCGATGAGTGT----- 839
 QY 1103 IleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetPro 1122
 Db 840 -----GTCCCTTACCCC-----GGCTGCTGTCATGGCAGTTGTGTGGAGCCC 881
 QY 1123 -----CysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
 Db 882 TGGCAGTGCAC-----TGT-----GAGCCACTGGGGCGGCGCTGCTGT 923
 QY 1141 LeuAlaCysProPheTyrGlyThrThrPropheAlaGlySerArgSerIleThrGluCys 1160
 Db 924 -----GACAAAGACCTGAACCTACTGT 944
 QY 1161 SerSerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeu 1180
 Db 945 GGCAGC----- 950
 QY 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPheAsn 1200
 Db 951 -----CAACCAC 956
 QY 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGly---TyrValCysLeuCys 1219
 Db 957 CCTGTCACCAACGGAGGCACGTCATGAACGCCGAGCGCTGACAGTACCGCTGCACCTGC 1016
 QY 1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
 Db 1017 CCTGACGGCTACTCGGCGCAGGAACCTGTGAGAAGGCTGACACCGCTGCACCTCCAAACCG 1076
 QY 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
 Db 1077 TGTGCAACGGGGCTCTTGCCATGAGGTGCGCTCGCGCTTCGATGCCACCTGCCCATCG 1136
 QY 1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeu 1279

1137 GGCTGGAGCGGGCCACCTGTCCTTGCATCGATGAGTGTCTTGCAGACCCGCTGTGCG 1196
1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
1197 GCGGTGGCACCTGTGTGGACGAGTGGACGGCTTTGAGTGCATCTGCCCGGAGCAGTGG 1256
1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
1257 GTGGGGGCCACCTGCCAGCTGGAGCGCCATGATGTGAGGAGGAGCCATGCCCTTAAGCT 1316
1320 AlaValCysGluAspGlnValGlyPheLeuCysLysCysProGlyPheLeuGly 1339
1317 TTTTCTTGCAGAAACCTGATTTGGCGCTATTACTGTGATTTGCATCCCGGGCTGGAAGGC 1376
1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValThr 1359
1377 ATCAACTGCCCATATCAACGTCACAGACTGTGCGGGCAG---TGTCAGCATGGGGCACC 1433
1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
1434 TGCAGGACCTGGTGAACCGGTACCGTGTGTGTGCCACCGGGCTTCGAGAGCGGCAT 1493
1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
1494 TGGAGCTGGAACGACGACAAAGTGTGCCAGCAGCCCTGCCACAGCGCGCCCTCTGCGAG 1553
1400 AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlnArgCysGlu 1419
1554 GACCTGGCGGACGGCTTCCACTGCCCTGCCCGGCGCTTCTCCGGGCTCTCTGTGAG 1613
1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
1614 GTGGAT-----GTC 1622
1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCys-----Thr 1454
1623 GACCTTTGTGAGCAAGCCCTGCGGAGCGCGCTGCTGCTATACCTGGAGGGTGAC 1682
1455 PheTyrMetLysSerSerAspMet-----AsnTyrGlyThrProLysTyr 1471
1683 TATTACTGGCGCTGCCCTGTGATGACTTTGGTGGCAAGAACTGCTCCGTCGCC 1733
1472 AlaValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeu 1491
1733 ----- 1733
1492 TyrValAsnGlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrPheHis 1511
1734 -----CGCGAGCG-----TGCCCT----- 1748
1512 HisIleAlaIleThrTyrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLys 1531
1749 -----GGCGGGCGCTGCGAGTG---ATCGATGGCTGC 1778
1532 LeuSerAspGlyValAlaGlyLeuSerValGlyLeuProIleProGlyGlyAlaLeu 1551
1779 GGTTCAGACCGGGCGCTGGAGT-----CTGGCAGCAGCGCTCC 1820
1552 ValLeu-----GlyGlnGluInAspLysLysGlyGluGlyPheSerProAla 1567
1821 GCGGTGTGTGGCCCCCATGGAGCTGCTGCGCAGCCAGCGGGGCAACTTTTCTGTCATC 1880
1568 -----GluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAspTyrValLeuSer 1585
1881 TGTGACAGTGGCTTTACTGCGACCTACTGCTCATGAGAACATTCAGCACTGCCCTGGGCCAG 1940
1586 Pro-----GlnGlnValLysSerLeuAlaThrSerCysPro 1597
1941 CCTTGGCGCAATGGGGGCATCATCATGATGAGGTGAGCGCTTCCGCTCTCTGCCCC 2000
1598 GluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGly 1617
2001 -----AGCGCTGGAGGGC 2015

1618 LysValLysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySer 1637
2016 GAGCTC---TGCAGACCAATCCACAGACTGCTTCCGATCCC----- 2057
1638 ValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPhe 1657
2058 ---TGCCACAGCGCGCGCTGTACAGACTG-----GTCAATGACTTC 2099
1658 CysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGln 1677
2100 -----TACTGTGCTGGCAGCAGCGC 2120
1678 TrpThrGlnProLeuProHis-----CysGluArgIleSerCysGlyVal 1692
2121 TGAAGGGCAGACCTGCCACTCCAGTTCAGTGGAGTTCAGTGGATGCTACACCTGC----- 2174
1693 ProProLeuGluAsnGly-----PheHisSerAlaAspPhe----- 1706
2175 -----AGCAAGGTGGCACCTGTACGACAGCGGCGACACCTTCGGCTGCGCC 2222
1707 -----TyrAlaGlySerThr-----ValThr 1713
2223 TGGCCCCCGCTGGAAGGGCAGCACCTGCGCGCTGCGCAAGACAGCAGCTGCGTGCCT 2282
1714 TyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe----- 1729
2283 AACCCCTGTGTGAATGTGGACCTGCTGGGCGCAGCGGGGCTCTCTCTCGCATCTGC 2342
1730 -----CysThrAspAsnGlySer----- 1735
2343 CGGAGCGCTGGGAGGGTGTGTACTTGTCACTCACAATACCAACGACTGCAACCCCTCTGCCT 2402
1736 ---TrpAsnGly----- 1738
2403 TGTACAAATGTGGCATCTGTGTGACGGGTCAACTGGTTCGCTGGAGTGTGCACCT 2462
1739 -----ValSerProSerCys---LeuAspValAspGluCysAlaValGlySerAspCys 1755
2463 GCGTTCGGGGGCTGACTGCGCATCAATCAGCAGTGCCAG---TCTTGGCCCTGT 2519
1756 SerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysValProPro 1775
2520 GCCTACGGGCGCAGCTGTGTGGATGAGATCAACGGGTATGCTGTAGCTGCCACCCCGC 2579
1776 TyrThrGlyAspGlyLysAsnCysAlaGluProIle-----LysCysLysAla 1791
2580 CGAGCC-----GGCCCCCGTGGCAGGAAGTGTATGCGGGAGATCTCTGTGTGCTCC 2633
1792 ProGlyAsnPro 1795
2634 CGGGGCACTCCG 2645

RESULT 20

US-09-855-722-4
; Sequence 4, Application US/09855722
; Patent No. 6638741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:


```
1586 Pro-----GlnGlnVallySerLeuAlaThrSerCysPro 1597
1941 CCCTGCCCAATGGGGGCACATGATGAGTGGAGCGCTTCGGCTGCTTGGCCC 2000
1598 GluGluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGly 1617
2001 -----AGCGGCTGGGAGGCG 2015
1618 LysVallyIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySer 1637
2016 GAGCTC---TGCAGACCAATCCCAACGACTGCCTTCCGATCCC----- 2057
1638 ValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPhe 1657
2058 ---TGCACAGCGCGCGCTGCTACGACTG-----GTCAATGACTTC 2099
1658 CysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGln 1677
2100 -----TACTGTGGCTGGCAGACGCG 2120
1678 TrpThrGlnProLeuProHis-----CysGluArgIleSerCysGlyVal 1692
2121 TGAAGGGCAAGACTGCTCACTCAGCGAGTTCCAGTGGGAGCTTACACCTGC----- 2174
1693 ProProLeuGluAsnGly-----PheHisSerAlaAspAspPhe----- 1706
2175 -----AGCAAGGTGGCACTGCTACGACAGCGGAGCACCTTCGGCTGGCGCC 2222
1707 -----TyralaGlySerThr-----ValThr 1713
2223 TGGCCCCCGCTGGAGGGCGAGCACCTGCGCGCTGCGCAAGAACAGCAGCTGCTGCGCC 2282
1714 TyrGlnCysAsnAngGlyTyrTyrLeuLeuGlyAspSerArgMetPhe----- 1729
2283 AACCCCTGTGTGAANGTGGCACCTGCTGGGCGCGCGGCGCTCTCTCTGCAATGCG 2342
1730 -----CysThrAspAngGlySer----- 1735
2343 CGGACGCGCTGGAGGGTGGTACTTGGCACTCAACATCAACAGCTGCAACCCCTGCGCT 2402
1736 ---TrpAngly----- 1738
2403 TGCTACAATGTTGGCATCTGTGTGACGGCTCAACTGTTCCGCTGCGAGTGTGCACCT 2462
1739 -----ValSerProSerCys---LeuAspValAspGluCysAlaValGlySerAspCys 1755
2463 GGCTTCGCGGGCGCTGACTGCGGCATCAACATCGACGAGTGGCCAG---TCTCGCCCTGT 2519
1756 SerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysValProPro 1775
2520 GCCTACGGGGCCAGTGTGTGATGATGATCAACGGGTATCGCTGTAGCTGCCACCGCGC 2579
1776 TyrThrGlyAspGlyLysAsnCysAlaGluProLe-----LysCysLysAla 1791
2580 CGAGCC-----GGCCCCCGGTGGCAGGAGTATCGGGTTCGGGAGATCTCTGTGTGTC 2633
1792 ProGlyAsnPro 1795
2634 CGGGGCACTCCG 2645
```

RESULT 21
US-08-400-159-7
; Sequence 7, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowitz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.

```
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE  
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/400,159  
; FILING DATE: 07-MAR-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-029  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4464 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; US-08-400-159-7  
  
Alignment Scores:  
Pred. No.: 1,67e-50 Length: 4464  
Score: 671.00 Matches: 233  
Percent Similarity: 34.38% Conservative: 86  
Best Local Similarity: 25.11% Mismatches: 271  
Query Match: 6.57% Indels: 338  
DB: 2 Gaps: 46  
  
US-09-977-053-6 (1-1842) x US-08-400-159-7 (1-4464)  
Qy 967 LeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986  
Db CTGGAGCTGCAGATCCGCGTGGCGAGAGAACTACTACAGCGCCACTTGCACACAG 448  
Qy PheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThr 1006  
Db TTCGCGCGCGCCGCAATGACTTTTCGCG----- 478  
Qy TyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGlu 1026  
Db -----CACTACACCTGC-----GACCAG 496  
Qy GluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSer 1046  
Db TACGGCAAC-----AAGGCTGCATGGACGCGCTGGATGGCAAG----- 535  
Qy ArgAsnIleSerAspCysLys---AlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065  
Db -----GAGTGCACAGAGCTGTGTGTAAACAGAGGGTGTAAATTGCTCCACGGG 583  
Qy LeuGluThrCysGluSerCysProLeu---GlyThrTyrGlnProLysPheGly----- 1082  
Db -----GGATGCACCGCTGCTGGCGAGTGCAGGTGCAGCTACGCTGGCAA 628  
Qy SerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsn 1102  
Db GGGAGTTCTCGCATGAGTGT----- 649
```

QY 1103 IleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetPro 1122
DB 650 -----GTCCCTTACCCC-----GGCTCGGTGCATGGCAGTGTGTGGAGCCC 691
QY 1123 -----CysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
DB 692 TGGCAGTGCAC-----TGT-----GAGACCACTGGGGCGGCTGCTGTGT 733
QY 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
DB 734 -----GACAAAGACCTGAATCTACTGT 754
QY 1161 SerSerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeu 1180
DB 755 GGCAGC----- 760
QY 1181 GlyHisIleLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
DB 761 -----CACAC 766
QY 1201 ProCysHisAsnSerGlyThrCys-----GlnGlnLeuGlyArgGlyTyr 1215
DB 767 CCTCTGACCAACGGAGGACGGTGCATCAACGCCGAGCCTGACCAG-----TAC 814
QY 1216 ValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCys 1235
DB 815 CGCTGCACCTGCCCTACTCGGCGCAAGAACTGTGAGAAGCGCTGAGCACGCCCTGC 874
QY 1236 SerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyLysPheIleCys 1255
DB 875 ACCTCAACCCGTGTGCCAACGGGGGCTCTTCCATGAGGTGCCGTCCGGCTTCGAATGC 934
QY 1256 GluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1275
DB 935 CACTGCCCATCGGCTGGAGCGGCCACCTGTGCCCTTGACATCGATGAGTGTGCTCG 994
QY 1276 SerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCys 1295
DB 995 AACCGGTGTGCCCGCGTGGCCACTGTGTGGACCAAGTGGAGTGGCGCTTTGAGTGCATCTGC 1054
QY 1296 ValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnPro 1315
DB 1055 CCGAGAGTGGTGGGGCCACTGCCAGCTGGAGCCCAATGATGTGAAGGAGCA 1114
QY 1316 CysLeuAsnAsnAlaValCysGluAspGlnValGlyPheLeuCysLysCysProPro 1335
DB 1115 TGCCCTTAACGCTTTTCTTGCACAAACCTGATGGCGCTATTACTGTGATTCATCCCG 1174
QY 1336 GlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLys 1355
DB 1175 GGCCTGAAGGGCATCACTGCCATATCAACGTCAACGACTGTGCGGGCAG-----TGTGAG 1231
QY 1356 AsnGlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPhe 1375
DB 1232 CATGGGGCACCTGCAGAGACCTGTGTGAACGGGTACCACTGTGTGCCCCAGGGGCTTC 1291
QY 1376 ThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGln 1395
DB 1292 GGAGGCGCGCATTCGAGCTGGAACGAGACAAAGTGTGCCAGCAGCCCTGCCACAGCGGC 1351
QY 1396 AlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGly 1415
DB 1352 GGCCTCTCGAGGACCTGGCCGACGCGCTTCCACTGCCACTGCCCCCAGGGCTTCTCCGGG 1411
QY 1416 GlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIle 1435
DB 1412 CCTCTCTGTGAGGTGAT----- 1429
QY 1436 TyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCys----- 1453
DB 1430 -----GTGACCTTTGTGAGCCACGCCCTTCCCGGAACGGCGCTCGCTGTCTATAAC 1480

QY 1454 -----ThrPheTyrMetLysSerSerAspAspMet-----AsnTyrGlyThr 1467
DB 1481 CTGGAGGGTGAATATCTATCTGCGCCTGCCCTGATGACTTTGGTGGCAGAACTGCTCCGTG 1540
QY 1468 ProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsn 1487
DB 1541 CCC----- 1543
QY 1488 GlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysProSerValAsnAsp 1507
DB 1544 -----CGCAGCGC-----TGCCCT----- 1558
QY 1508 GlyArgTyrHisHisIleAlaIleThrTyrThrSerAlaAsnGlyIleTyrLysValTyr 1527
DB 1559 -----GGCGGGCGCTCCAGAGTG----- 1576
QY 1528 IleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
DB 1577 ATCGATGGCTGGGGTTCAGACGGGGCTCGGATG-----CCTGGC 1618
QY 1548 GlyGlyAlaLeuValLeu-----GlyGlnGluGlnAspLysGlyGluGly 1563
DB 1619 ACAGCAGCCTCGGGGTGTGGCCCCCATGAGCGCTGCGTACGCCAGCCAGGGGCAAC 1678
QY 1564 PheSerProAla-----GluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAsp 1581
DB 1679 TTTTCTGCATCTGTGACAGTGGCTTTACTGGCACCCTACTGCCATGAGAACTTGGACAC 1738
QY 1582 TyrValLeuSerPro-----GlnGlnValLysSerLeuAla 1593
DB 1739 TGCTGGGGCAGCCCTCGCGCAATGGGGGCACATGATCGATGAGGTGAGCGCTTCCGC 1798
QY 1594 ThrSerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSer 1613
DB 1799 TGCTTCTGCCCC-----AGC 1813
QY 1614 GlyIleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg 1633
DB 1814 GFTTGGAGGGCGAGCTC-----TGCGACACCAATCCCAACGACTGCTTCCCGATCCC--- 1867
QY 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
DB 1868 -----TGCCACAGCGCGCGCGCTGTCTACGACCTG----- 1897
QY 1654 ValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeu 1673
DB 1898 GTCAATGACTTC-----TACTGTGCG 1918
QY 1674 AsnGlnGlyGlnTyrThrGlnProLeuProHis-----CysGluArgIle 1688
DB 1919 TGCGAGCAGCGCTGGAAGGCAAGACCTGCCACTCAGCGAGTTCAGTGCAGTGCCTAC 1978
QY 1689 SerCysGlyValProProProLeuGluAsnGly-----PheHisSerAlaAspAsp 1705
DB 1979 ACCTGC-----AGCAACGGTGGCACCCTGCTACGACAGCGCGCACC 2020
QY 1706 Phe-----TyrAlaGlySerThr----- 1711
DB 2021 TTCCGCTGCGCTGCCCCCCCCCGCTGGAAGGGCAGCAGCCTTGCCCGCTGCGCAAGACAGC 2080
QY 1712 -----ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe 1729
DB 2081 AGCTGCTGCCCAACCCCTGTGTGATGTGGCACCCTGCTGGCGGAGCGGGCTCTCCTTC 2140
QY 1730 -----CysThrAspAsnGlySer----- 1735
DB 2141 TCCTGCATCTGCGCGGACCGCTGGAGGGTGTGTTGCACTTGTCACTCAATACCAAGACTGC 2200
QY 1736 -----TrpAsnGly----- 1738
DB 2201 AACCTCTGCTGTGTGATGATGTCATCTGTGTGACGGCGTCAACTGCTTCCGCTGC 2260
QY 1739 -----ValSerProSerCys---LeuAspValAspGluCysAlaVal 1751

Db 2261 GAGTGTGACCTGGCTTCGGGGCCCTGACTGCGCATCAACATGACGAGTGCAG--- 2317
 QY 1752 GlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerThrLeuCysSer 1771
 Db 2318 TCCTGCGCCCTGTGCTACGCGGGCCAGCTGTGTGATGAGATCAACGGGTATCGCTGTAGC 2377
 QY 1772 CysValProProThrGlyAspGlyLysAsnCysAlaGluProIle--- 1787
 Db 2378 TGCCCAACCGCGCGAGCC-----GGCCCGGTGCCAGGAGTGTGCGGTTCGGGAGA 2431
 QY 1788 LysCysLysAlaProGlyAsnPro 1795
 Db 2432 TCCTGTGTGCTCCGGGGCACTCG 2455

RESULT 22

US-08-611-729A-7
 ; Sequence 7, Application US/08611729A
 ; Patent No. 6004924
 ; GENERAL INFORMATION:
 ; APPLICANT: Ish-Horowicz, David
 ; APPLICANT: Henrique, Domingos M.P.
 ; APPLICANT: Lewis, Julian H.
 ; APPLICANT: Myat, Anna M.
 ; APPLICANT: Fleming, Robert J.
 ; APPLICANT: Artavanis-Tsakonas, Spyridon
 ; APPLICANT: Mann, Robert S.
 ; APPLICANT: Gray, Grace B.
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
 ; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/611,729A
 ; FILING DATE: 06-MAR-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Miarock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-037
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 4483 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 332..4102
 US-08-611-729A-7

Alignment Scores:
 Pred. No.: 1.68e-50
 Score: 671.00
 Percent Similarity: 34.38%
 Best Local Similarity: 25.11%
 Query Match: 6.57%
 Length: 4483
 Matches: 233
 Conservative: 86
 Mismatches: 271
 Indels: 338

DB: 3 Gaps: 46
 US-09-977-053-6 (1-1842) x US-08-611-729A-7 (1-4483)
 QY 967 LeuAlaSerGluLeuLeuLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986
 Db 389 CTGAGCTGCAGATCGCGTGCAGACGAGAACTACTACACGCCCACTTGCACCAAG 448
 QY 987 PheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThr 1006
 Db 449 TTCTGCGCGCCCGCAATGACTTTTTCGGC----- 478
 QY 1007 TyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGlu 1026
 Db 479 -----CACTACACTTGC-----GACCAG 496
 QY 1027 GluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSer 1046
 Db 497 TACGGCAAC-----AAGGCTGCATGCGCGCTGGATGGGCAAG----- 535
 QY 1047 ArgAsnIleSerAspCysLys---AlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065
 Db 536 -----GAGTGCACAGGAAGCTGTGTGTAACACAGGGTGTAAATTTGCTCCACGGG 583
 QY 1066 LeuGluThrCysGluSerCysProLeu---GlyThrTyrGlnProLysPheGly----- 1082
 Db 584 -----GGATGCACCGTGCCTGGGGAGTGCAGGTGCAGCTACGGCTGGCA 628
 QY 1083 SerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsn 1102
 Db 629 GCGAGGTTCTCGATGAGTGT----- 649
 QY 1103 IleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetPro 1122
 Db 650 -----GTCCCTTACCCC---GGCTGCTGCATGCGAGTGTGTGGAGCCC 691
 QY 1123 -----CysHisProCysProArgAspTyrTyrGlnProAlaGlyLysAlaPheCys 1140
 Db 692 TGCAGTGCAC---TGT-----GAGACCACTGGGGGGGCTGCTCTGT 733
 QY 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
 Db 734 -----GACAAAGACCTGACTACTGT 754
 QY 1161 SerSerPheSerSerThrPheSerAlaAlaGluSerValValProAlaSerLeu 1180
 Db 755 GGCAGC----- 760
 QY 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
 Db 761 -----CAACCAC 766
 QY 1201 ProCysHisAsnSerGlyThrCys-----GlnGlnLeuGlyArgGlyTyr 1215
 Db 767 CCTTGCACCAACGAGGACGATCATCAACGCCGAGCTGACACAG-----TAC 814
 QY 1216 ValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCys 1235
 Db 815 CGCTGCACCTGCGCTGACCGCTACTCGGCGAGGAACCTGTGAGAGGCTGACACGCGCTGC 874
 QY 1236 SerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCys 1255
 Db 875 ACCTCCAACCGGTGTGCCAACGGGGGCTCTTGCATGAGGTGCGCTCGGGTTCGATGC 934
 QY 1256 GluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1275
 Db 935 CACTGCGCATCGGGCTGGAGCGGCGCCACCTGTGCTCCCTTGACATCGATGAGTGTCTCG 994
 QY 1276 SerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCys 1295
 Db 995 AACCCGTGTGCGCGCGCTGGACCTGTGTGGACAGGTGACCGCTTGTAGTGCATCTGC 1054
 QY 1296 ValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnPro 1315

Db 1055 CCCGAGCAGTGGGTGGGGCCACCTGCCAGCTGGAGCCCAATGAGTGAAGGGAGGCA 1114
Qy 1316 CysLeuAsnAlaValCysLeuAspGlnValGlyGlyLeuCysLeuCysPro 1335
Db 1115 TGCCTTAACGCTTTCTTGAACAACTGATGGGGCTATTACTGTGATGATCCG 1174
Qy 1336 GlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLys 1355
Db 1175 GCCTGGAAGGGCATCAACTGTCATCAACGCTCAACGACTGTGCGGGCAG---TGTGAG 1231
Qy 1356 AsnGlyAlaThrCysLeuAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPhe 1375
Db 1232 CATGGGGGACCTGCAAGGACCTGTGTGAACGGGTACCAAGTGTGTGTGCTCCAGGGGCTTC 1291
Qy 1376 ThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGln 1395
Db 1292 GGAGGCGGCGATTCGAGCTGGAACGAGACAGTGTGCCAGCAGCCCTGTCACAGCGGC 1351
Qy 1396 AlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGly 1415
Db 1352 GGCCTTCGAGGAGCTGGCGACGGCTTCCACTGCCACTGCCCGCAGGGCTTCTCCGGG 1411
Qy 1416 GlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIle 1435
Db 1412 CCTCTCTGTGAGGTGGAT----- 1429
Qy 1436 TyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCys 1453
Db 1430 -----GTGACCTTTGTGAGCAAGCCCTGCCGAGCGGCTGCTGTATTAAC 1480
Qy 1454 -----ThrPheTrpMetLysSerSerAspMet-----AsnTyrGlyThr 1467
Db 1481 CTGAGGGTGACTATTACTGCGCTGCTGATGACTTTGGTGGCAAGAACTGCTCCGTG 1540
Qy 1468 ProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuThrAspTyrAsn 1487
Db 1541 CCC----- 1543
Qy 1488 GlyTrpValLeuTyrValAsnGlyArgGluLysIleThrAsnCysProSerValAsnAsp 1507
Db 1544 -----CGCAGCGG-----TGCCCT----- 1558
Qy 1508 GlyArgTrpHisIleAlaIleThrTrpThrSerAlaAsnGlyIleTrpLysValTyr 1527
Db 1559 -----GGCGGGGCTGCGAGAGT----- 1576
Qy 1528 IleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
Db 1577 ATCGATGCTGGGTGAGCGCGGGGCTGAGT-----CCTGCG 1618
Qy 1548 GlyGlyAlaLeuValLeu-----GlyGlnGluGlnAspLysGlyGly 1563
Db 1619 ACAGCAGCTCGGGGTGTGGCCCATGAGCGCTGCTGAGCCAGCCAGGGGGCAAC 1678
Qy 1564 PheSerProAla-----GluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAsp 1581
Db 1679 TTTTCTGTCATCTGTGACAGTGGCTTTACTGTCACCTACTGTCATGAGAACATTGACGAC 1738
Qy 1582 TyrValLeuSerPro-----GlnGlnValLysSerLeuAla 1593
Db 1739 TGCTGGCCAGCCCTGCGCAATGGGGGCAATGATGATGAGGTGGAGCCCTTCGCG 1798
Qy 1594 ThrSerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSer 1613
Db 1799 TGCTTCTGCCCC-----AGC 1813
Qy 1614 GlyIleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg 1633
Db 1814 GTTGGGAGGGCGAGCTC---TGCGACCAACCAATCCACGACTGCTTCCGATCCC--- 1867
Qy 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653

Db 1868 -----TGCCACAGCCGCGGCGCTGCTAGCACCTG----- 1897
Qy 1654 ValLeuLeuPheCysAspProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeu 1673
Db 1898 GTCAATGACTTC-----TACTGTGCG 1918
Qy 1674 AsnGlnGlyGlnTrpThrGlnProLeuProHis-----CysGluArgIle 1688
Db 1919 TGCGACGAGCTGGAAGGCAAGACCTGCCACTCAOCGAGTTCAGTGCAGTGCCTAC 1978
Qy 1689 SerCysGlyValProProLeuGluAsnGly-----PheHisSerAlaAspAsp 1705
Db 1979 ACTTGC-----AGCAACGCTGGCACCCTGCTACGACAGCGCGACACC 2020
Qy 1706 Phe-----TyrAlaGlySerThr----- 1711
Db 2021 TTCCGCTGCGCTGCCCGGCTGGAGGGGAGCAGCACCCTGCGCGCTGCCAAGACAGC 2080
Qy 1712 -----ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe 1729
Db 2081 AGTGCCTGCCAACCCCTGTGTGAATGGTGGCACCCTGCTGGGCGAGCGGGGCTCTTC 2140
Qy 1730 -----CysThrAspAsnGlySer----- 1735
Db 2141 TCCTGCATCTGCGGAGCGCTGGGAGGGTGTACTTGCCTCACAATACCAAGACTGC 2200
Qy 1736 -----TrpAsnGly----- 1738
Db 2201 AACCTCTGCTTGTACTACATGTGTGGCATCTGTGTGAGCGGCTCAACTGGTTCCGCTGC 2260
Qy 1739 -----ValSerProSerCys-----LeuAspValAspGluCysAlaVal 1751
Db 2261 GAGTGTGACCTTGTGCGGGGCTGACTCCGCTCATCACTCCAGTGCAGTGCAG--- 2317
Qy 1752 GlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSer 1771
Db 2318 TCCTGCGCTGTGCTACGGGCGCAGTGTGGATGAGATCAACGGGTATCGCTGTAGC 2377
Qy 1772 CysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIle----- 1787
Db 2378 TGCCACCGCGCGAGCC-----GGCCCCCGGTGCCAGAGTGTATCGGTTCGGGAGA 2431
Qy 1788 LysCysValysAlaProGlyAsnPro 1795
Db 2432 TCCTGCTGCTCCGCGGCACTCCG 2455

RESULT 23

US-08-882-046-3
Sequence 3, Application US/08882046
Patent No. 6136952

GENERAL INFORMATION:

APPLICANT: Li, Linheng
APPLICANT: Hood, Leroy
APPLICANT: Krantz, Ian D.
APPLICANT: Spinner, Nancy B.
TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
TITLE OF INVENTION: Nucleic Acids and Methods of Use
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/882,046
FILING DATE: 25-JUN-1997

CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-UW 2637
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4315 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 16..3460
 US-08-882-046-3

Alignment Scores:
 Pred. No.: 6,47e-48 Length: 4315
 Score: 642.50 Matches: 232
 Percent Similarity: 33.68% Conservative: 93
 Best Local Similarity: 24.04% Mismatches: 281
 Query Match: 6.29% Indels: 360
 DB: Gaps: 45

US-09-977-053-6 (1-1842) x US-08-882-046-3 (1-4315)

QY 967 LeuAlaSerGluIleLeuLeuAlaAspSerAsnSerLeuGluThrIleLysAlaSerPro 986
 DB 433 CTGGAGCTCAGATCGCGCGCTCGACGAGAACTACTACAGCGCACTTGCACACAG 492
 QY 987 PheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThr 1006
 DB 493 TTCTGCGCGCCCGCAACGACTTTTTCGGC----- 522
 QY 1007 TyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGlu 1026
 DB 523 -----CACTACACCTGC-----GACCAG 540
 QY 1027 GluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSer 1046
 DB 541 TACGGCAAC-----AAGGCCTGCATGCGACGGCTGGATGGGCAAG----- 579
 QY 1047 ArgAsnIleSerAspCysLys-----AlaGlnCysLysGlnGly----- 1059
 DB 580 -----GAGTGCACGAAGCTGTGTGTAAACAAGGGGTAAATTTGCTCCACGGG 627
 QY 1060 -----ThrTyrSerTyrSerGlyLeuGluThrCys 1069
 DB 628 GGATGACCGCTCGCTGGGAGTGCGAGTGCGAGCTACGCTGCGCAAGGG-----AGGTTCTGC 684
 QY 1070 GluSerCysProLeuGlyThrTyrGlnProLysPheGlySerArgSerCysLeu----- 1087
 DB 685 GATGAGTGT-----GTCCCTACCCCGGC-----TGCGTGCATGCG 720
 QY 1088 SerCysProGlu-----AsnThrSerThrValLysArgGlyAlaVal----- 1101
 DB 721 AGTTGTGTGGAGCCCTGGCGAGTGCGCACTGTGAGACCAACTGGGGCGCGCTGTCTGTGAC 780
 QY 1102 ---AsnIleSerAlaCysGlyVal-----ProCysProGluGlyLysPheSerArgSer 1118
 DB 781 AAGACCTGAAGTACTGTGGGAGCCACACCTGACCAACGGA-----GGC 828
 QY 1119 GlyLeuMetProCysHisPro-----CysProArgAspTyrTyrGln 1132
 DB 829 ACGTGCATCAACCGCGAGCTTACCATCGCTGACCTGCGCTGACCGGCTACTCGGCG 888
 QY 1133 ProAsnAlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAla 1152
 DB 889 AGGAACCTGTGAGAGGCT----- 906

QY 1153 GlySerArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGlu 1172
 DB 906 ----- 906
 QY 1173 SerValProProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGln 1192
 DB 906 ----- 906
 QY 1193 ValPheHisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnLeuGly 1212
 DB 907 ---GAGCACCGCTGACCTCAACCGCTGGCCAAACGGGGCTCTTGGCATGAGCGCG 963
 QY 1213 ArgGlyTyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIle 1232
 DB 964 TCGGCTTCGANTGCCACTGCCCATCGGCTGGAGCGGGCCACCTGTGCCCTTGACATC 1023
 QY 1233 AspGluCysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGlu 1252
 DB 1024 GATGAGTGTCTTCAACCGCTGTGGCGCGGTGGCCACCTGTGTGGACACGAGTGCACGC 1083
 QY 1253 PheIleCysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGlu 1272
 DB 1084 TTTGAGTGCATCTGCCCGGAGCAGTGGGTGGGGGCCACCTGCCAGCTGAGCGTCAACAG 1143
 QY 1273 CysSerSerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArg 1292
 DB 1144 TGTGCGGGCGAG---TGTACATGGGGGCACCTGCAAGNACCTGTGTGAACGGGTACCAG 1200
 QY 1293 CysThrCysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGln 1312
 DB 1201 TGTGTGTGCCACCGGGCTTTCGGAGGCGCGCATTCGCGAGCTGGAAACGACAGTGTGCC 1260
 QY 1313 SerAsnProCysLeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLys 1332
 DB 1261 AGACGCCCCGTGCCACAGCGCGGCTCTGCGAGGACCTGCGCGACGGCTTCACATGCCAC 1320
 QY 1333 CysProProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGln 1352
 DB 1321 TGCCCCCAGGGCTTCTCCGGGCGCTCTCTGTGAGGTGATGTGACCTTTGTGAGCCAGC 1380
 QY 1353 ProCysLysAsnGlyAlaThrCys----- 1360
 DB 1381 CCTTCCGGAACCGCGCTGCTGTATACTTGGAGGTGACTATTACTTGGCGCTGCCCT 1440
 QY 1360 ----- 1360
 DB 1441 GATGACTTTGGTGGCAAGAACTCTCGTGCCTCCCGAGCGGTGCTGCGGGGCGCTGCAG 1500
 QY 1361 -----LysAspGlyAla----- 1364
 DB 1501 AGTGATCGATGGCTGGGCTCAGACCGGGGCTTGGGATGCTGGCACGACGCTCCGGC 1560
 QY 1365 -----AsnSerPheArgCysLeuCys 1371
 DB 1561 GTGTGTGGCCCCCATGAGCGCTGGTCCAGCCAGCGGGCAACTTTTCTGTCATCTGT 1620
 QY 1372 AlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnPro 1391
 DB 1621 GACAGTGGCTTTTACTGGCACCTACTGTCATGAGTGGAGCGCTTCCGCTGCTTCTGCCCCAGC 1680
 QY 1392 CysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnPro 1411
 DB 1681 TGCGCAATGGGGGCACATGATCGATGAGTGGAGCGCTTCCGCTGCTTCTGCCCCAGC 1740
 QY 1412 GlyPheSerGlyGlnArgCysGluThrGlu-GlnSerThrGlyPheAsnLeuAspPheGln 1431
 DB 1741 GGCTGGAGGCGCGAGCTGTGCGACACCAATCCCAACAGACTGCCTTC----- 1786
 QY 1431 uValSerGlyIleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLe 1451
 DB 1787 -----CCGATCCCTGCCACAGCGCGCGC 1809
 QY 1451 uThrCysThrPheTrpMetLysSerSer-----AspAspMetAsnTyrGlyThr 1467

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 10, 2004, 15:48:01 ; Search time 2199.65 Seconds
(without alignments)
3793.447 Million cell updates/sec

Title: US-09-977-053-6

Perfect score: 10219

Sequence: 1 MWPRLLAFCCGGLALVSGWAT.....GVTKITCLSGWNLIPYC 1842

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09977053/runat 06052004 075944 18296/app query.fasta_1.5710
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=25 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09977053@cgn_1_484@runat 06052004 075944 18296
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	10208	99.9	10878	9	US-09-911-842-1
2	10208	99.9	10878	14	US-10-150-821-1
3	10047	98.3	10674	16	US-10-295-027-429
4	10047	98.3	10674	16	US-10-295-027-1080
5	9862	96.5	11152	16	US-10-028-248A-7
6	9862	96.5	11152	16	US-10-107-782-7
7	9812	96.0	11158	16	US-10-028-248A-5
8	9812	96.0	11158	16	US-10-107-782-5
9	8522.5	84.7	11230	9	US-09-911-842-3
10	8522.5	84.7	11230	14	US-10-150-821-3
11	5485	53.7	3448	13	US-10-302-172-34
12	3133	30.7	2064	10	US-09-822-846-55
13	1977	19.3	1408	9	US-09-822-846-54
14	1498	14.7	1892	9	US-09-764-898-49
15	1278	12.5	680	9	US-09-764-853-173
16	1278	12.5	680	9	US-09-764-898-122
17	1278	12.5	680	10	US-09-764-881-14
18	1278	12.5	680	13	US-09-764-881-14
19	1278	12.5	680	13	US-09-764-893-37
20	1278	12.5	680	15	US-10-073-865-37
21	1278	12.5	680	16	US-10-242-747-14
22	1237	12.1	3262	13	US-10-466-164-29
23	914.5	8.9	6728	16	US-10-190-115-3
24	914.5	8.9	6728	16	US-10-369-072-3
25	893.5	8.7	7410	16	US-10-190-115-1
26	893.5	8.7	7410	16	US-10-369-072-1
27	872	8.5	7596	15	US-10-004-113-57
28	872	8.5	8064	15	US-10-004-113-56
29	869.5	8.5	8221	12	US-10-152-319A-2167
30	869	8.5	7693	13	US-10-072-012-133
31	860	8.4	7673	9	US-09-815-925-1
32	860	8.4	7673	16	US-10-294-006-1
33	853	8.3	7332	13	US-09-944-849-1
34	853	8.3	7693	13	US-10-160-497-11
35	853	8.3	7693	13	US-10-348-750-11
36	837.5	8.2	2460	16	US-10-297-639-13
37	831.5	8.1	7615	15	US-10-004-113-59
38	829	8.1	8091	15	US-10-101-510-86
39	829	8.1	8091	15	US-10-356-625-1
40	829	8.1	8091	16	US-10-159-563-322
41	759	7.4	6122	17	US-10-272-810-3
42	759	7.4	6122	17	US-10-273-070-3
43	749	7.3	4208	9	US-09-855-722-6
44	749	7.3	4208	15	US-10-219-248-6
45	749	7.3	4208	15	US-10-219-247-6
46	745	7.3	5458	15	US-10-213-329-2
47	744	7.3	4208	9	US-09-995-593A-10
48	744	7.3	5590	13	US-10-236-392-55
49	744	7.3	5929	14	US-10-071-766-133
50	739.5	7.2	4146	15	US-10-303-685-3
51	739.5	7.2	4361	15	US-10-303-685-1
52	739.5	7.2	4413	13	US-10-243-552-783
53	739.5	7.2	4950	15	US-10-303-685-2
54	733	7.2	4200	16	US-10-093-463-137
55	727	7.1	3825	15	US-10-303-685-10
56	727	7.1	3921	15	US-10-303-685-7
57	724.5	7.1	3540	16	US-10-369-493-29658
58	721	7.1	4200	16	US-10-093-463-139
59	716.5	7.0	3786	15	US-10-303-685-8
60	714.5	7.0	2757	9	US-09-764-898-67
61	713	7.0	462	9	US-09-764-898-11
62	713	7.0	462	9	US-09-764-898-85
63	704	6.9	1713	15	US-10-133-013-101
64	691.5	6.8	5848	15	US-10-310-002-44
65	687.5	6.7	3027	15	US-10-310-002-8
66	687.5	6.7	4926	15	US-10-310-002-5
67	686.5	6.7	3612	15	US-10-303-685-9
68	674	6.6	3955	9	US-09-855-722-4
69	674	6.6	3955	15	US-10-219-248-4
70	674	6.6	3955	15	US-10-219-247-4
71	673	6.6	4333	13	US-10-236-392-57
72	672	6.6	4749	15	US-10-091-625-3

73 6.6 4749 15 US-10-096-399A-3 Sequence 3, Appli
74 6.6 4749 16 US-10-461-568-3 Sequence 3, Appli
75 6.6 4749 16 US-10-388-263-376 Sequence 376, App
76 6.6 4702 9 US-09-880-107-1545 Sequence 1545, Ap
77 6.6 4974 15 US-10-091-625-10 Sequence 10, Appl
78 6.6 4974 15 US-10-096-399A-10 Sequence 10, Appl
79 6.6 4974 16 US-10-461-668-10 Sequence 10, Appl
80 6.6 4974 16 US-10-388-263-380 Sequence 380, App
81 6.6 4974 16 US-10-294-006-23 Sequence 23, Appl
82 6.5 2631 16 US-10-115-479-79 Sequence 79, Appl
83 6.5 4401 16 US-10-310-002-7 Sequence 7, Appli
84 6.4 4524 15 US-10-310-002-7 Sequence 9, Appli
85 6.4 3087 15 US-10-310-002-7 Sequence 9, Appli
86 6.4 6331 15 US-10-310-002-45 Sequence 45, Appl
87 6.4 9662 16 US-10-159-563-286 Sequence 286, App
88 6.4 9728 15 US-10-084-817-35 Sequence 35, Appl
89 6.3 9835 15 US-10-198-846-13061 Sequence 13061, A
90 6.3 10166 16 US-10-295-027-161 Sequence 161, App
91 6.3 10172 9 US-09-954-456-702 Sequence 702, App
92 6.3 10172 10 US-09-960-706-937 Sequence 937, App
93 6.3 4413 16 US-10-115-479-81 Sequence 81, Appl
94 6.3 8942 16 US-10-191-803-223 Sequence 223, App
95 6.3 2325 10 US-09-900-449A-1 Sequence 1, Appli
96 6.3 10284 9 US-09-974-298-46 Sequence 46, Appl
97 6.3 10284 14 US-10-044-090-47 Sequence 47, Appl
98 6.2 2508 9 US-09-908-322-1 Sequence 1, Appli
99 6.2 2508 10 US-09-783-931-1 Sequence 1, Appli
100 6.2 2883 9 US-09-908-322-3 Sequence 3, Appli
100 6.2 2883 10 US-09-783-931-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-911-842-1
; Sequence 1, Application US/09911842
; Patent No. US20020151483A1
; GENERAL INFORMATION:
; APPLICANT: Welcher, Andrew A.
; APPLICANT: Elliott, Gary S.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911,842
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 10878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-911-842-1

Alignment Scores:
Pred. No.: 0 Length: 10878
Score: 10208.00 Matches: 1838
Percent Similarity: 99.95% Conservative: 3
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.89% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-911-842-1 (1-10878)

QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
DB 11 ATGTGGCTCGCCCTGGCCCTTTTGTGTGTGGGGTCTGGGGCTGGTTCTGGGGTGGGGGACC 70
QY 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
DB 71 TTTCAGCAGATGTCCCGTCCGCAATTCAGCTTCGCCCTTCCTCCCGCAGACGGCGCCC 130
QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60

DB 131 GGGGGCCCGGGAGTATCCCGCGCGCGCGCTCTGGCGACGACGAGCGGGAGCAGA 190
QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
DB 191 GTGAGCGGCTGGGCGCGCGCTTCGGCGAGCGGTGCGGTGCTGGGGAGCTCAGCGAG 250
QY 81 ArgLeuGluLeuValPheLeuValAspAspSerSerSerValGlyGluValAsnPheArg 100
DB 251 GCGCTGGAGCTGTCTCTCTGTGTGATGTTGTCCTGCTCCAGCTGGGGAGTCACTTCGCC 310
QY 101 SerGluLeuMetPheValArgGlyLeuLeuSerPhePheProValValProThrAlaThr 120
DB 311 AGCGAGCTCATGTTGTCGCGCAGCTGCTGTCGCTCCAGCTGGGGAGTCACTTCGCC 370
QY 121 ArgValAlaIleValThrPheSerSerIleValValProArgValAspTrpIle 140
DB 371 GCGGTGGCCATCGTACCTTCTGTCGCAAGAACCTACGTGGTGGCGCGGTGTGATACATC 430
QY 141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnLeuProAlaIle 160
DB 431 TCCACCGCGCGCGCGCGCAGCAGCAAGTGGCGCTGCTCTCCAGAGATCCCTGCCATC 490
QY 161 SerThrArgGlyGlyThrThrThrGlyAlaPheGlnAlaAlaGlnIleLeu 180
DB 491 TCTTACCGAGGTGGCGGACCTACCAAGCGGCTTCCAGCAAGCGCGCAAAATCTT 550
QY 181 LeuHisAlaArgGluLeuSerThrLysValValPheLeuIleThrAspGlyTrpSerAn 200
DB 551 CTTTCATGCTAGAGAAACTACAAAGTTGTATTTCTCATCATCTGATGGATATTCCAAT 610
QY 201 GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
DB 611 GGGGGAGACCTAGACCAATTGCGAGCTCATCGAGATTCAGGAGTGGAGATCTTCACT 670
QY 221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
DB 671 TTGGCATATGGCAAGGGAACATTGAGAGCTGAATGACATGGCTTCCACCCCAAGAG 730
QY 241 GluHisCysTrpLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260
DB 731 GAGCACTGTACTCTGCTACACAGTTTGTGAAGAAATTTAGGCTTTAGCTCCCGGGGATG 790
QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrpLeu 280
DB 791 CATGAAGATCTACTTCTGGAGTTTATTTCAAGATGATATGCTTCCACTGCTCTTATCTT 850
QY 281 CysAspGluGlyLeuAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
DB 851 TGTGATGAAGGCAAGAGCTGCTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACACA 910
QY 301 GlyHisPheGluCysIleCysGluLysGlyTrpTrpGlyLysGlyLeuGlnTrpGluCys 320
DB 911 GGCCATTTTGTAGTGCATCTGTGAAAGGGGTATTACGGGAAGGTCTGCAATGATGATGC 970
QY 321 ThrAlaCysProSerGlyThrTrpLysProGluGlySerProGlyGlyIleSerSerCys 340
DB 971 ACAGCTTGCCTACCTGGGGACATCAAACTGAAAGCTCACCAGGAGAAATCAGAGTTGC 1030
QY 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
DB 1031 ATTCCATGTCCTGATGAATAATCACACCTCTCCACTGGGAAGCACATCCCTGAAGACTGT 1090
QY 361 ValCysArgGluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
DB 1091 GTCTGCAGAGGGGATACAGGGCATCTGGCCAGACCTGTGAACTTGTTCACCTGCCCTGCC 1150
QY 381 LeuLysProProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
DB 1151 CTGAAGCCTCCGAAATGGTTACTTTATCCAAACACTTGCACCAACACCACTTCAATGCA 1210
QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
DB 1211 GCCTGTGGGTCCCATGTACCCCTGGATTGTATCTTGTGGGAAGCAGCATCATCTTATGT 1270

QY 421. LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
DB 1271 CTACCCAAATGGTTGTGGTCCGGTTACAGAGACTTACGACAGTAAGAAATATGTCCTCAT 1330
QY 441. LeuArgGlnProLysHisGlyHisLeuSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1331 CTCGGCAGCCGAACATNGCCACATCAGCTGTCTACAGGGGAAATGTTATATAGACA 1390
QY 461. ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1391 ACATGTTTGGTTCCTGTGTATGAAGGTCACAGACTAGAGGCGATGATAAGCTTACTTGT 1450
QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1451 CAAGGAACAGCCAGTGGGATGGGCCAGAACCCCGGTGTGTGAGGCCACTGTGTCCACC 1510
QY 501 PheGlnMetProLysAspValIleLeuSerProHisAsnCysGlyLysGlnProAlaLys 520
DB 1511 TTTTCAGATGCCAAGATATCAT 1570
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1571 TTTGGGACGATCTGCTATGTAAGTTGCGGCCAAGGGTTCATTTATCTGAGTCAGAA 1630
QY 541 MetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaValCysLys 560
DB 1631 ATGCTGAGATGTACCACTTCTGGAAATGGAATGTGCGAGTTGAGGCGCTGTGTGTAAA 1690
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGln 580
DB 1691 GAGCTGGAGGCTCTCAATCACTGTCTTAAGGACATAGAGGCTAAGACTCTGGAACAG 1750
QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 1751 CAAGATTCTGCCAATGTATCTGCGAGATTCCACAGCTAAGACAACTCTGTGTGAAAG 1810
QY 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
DB 1811 GTGTGAGTCACGTTTCATCCAGCTTTCACCCCACTTACCTTTTCCAGTTGGAGATGT 1870
QY 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
DB 1871 GCTATGCTATACGGCACTGACCTATTCGGCAACCGCCAGCTGCATTTCCATATC 1930
QY 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
DB 1931 AAGGTTATTGATGCAGAACCACTGTCATAGACTGGTGACAGATCTCCACCTCCGTCAG 1990
QY 661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
DB 1991 GTCTCGGAGAAGGTATACATGCCGCAAGCTGGGATGAGCCTCAGTTCTCAGACAACTCAGG 2050
QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
DB 2051 GCTGAATTGCTATTACAGAGTATACACAGAGAGACCTTTTCCCTCAAGGGGAGACT 2110
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
DB 2111 ATAGTACAGTATACAGCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATT 2170
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheLysCys 740
DB 2171 GTCATAAAGGTTTCTCCCTGTGAAATTCATATCACACCTGTAAATGGGATTTTATATGC 2230
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
DB 2231 ACTCCAGATATATCTGAGTCAACTGTACATTAATCTGTGGAGGCTATGATTTCACA 2290
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
DB 2291 GAAGGGTCTACTGACAAATATTATGTGCTTATGAAGATGGCGCTCGAAGAACCAATAT 2350

QY 781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
DB 2351 ACCACTGAATGGCAGACTGTGCCAAAAACGTTTTTCTAACACAGGGTTCAAGTCCTTT 2410
QY 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
DB 2411 GAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAAGAGTTTTCTGAA 2470
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
DB 2471 GCATTTGAGACGACCCCTGGGAAAAATGGTCCCATCTTTGTAGTGATGAGAGACATT 2530
QY 841 AspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
DB 2531 GACTGACAGACTGGAGGAGAACCTGACCAAAAAATATTGCTAGTAATATAATTATGACTAT 2590
QY 861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
DB 2591 GAAAAATGGCTTTGCATTTGACACAGGTGGCTGGGGTGCAGCTAATAGGCTGGAATTACTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGlnThrAlaThrSerIleGlyAsnAlaLysSer 900
DB 2651 TAGCATGACTTCTCGGACACTGTGCAAGAACAGCCCAAGCATCGGCAATGCCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
DB 2711 TCACGGATTAAAGAGAGTGGCCCATTTATCTGACTATTAATAATTAGTTAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
DB 2771 ACAGTAGTGTGCCATTACCCGATGAAAGAAATGATACCTTGAATGGGAAAAATCAGCA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
DB 2831 CGACTCCTTCAGACATTGGAACTATCACAAATAAACTGAAAAGAGACTCTCAACAAAGAC 2890
QY 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
DB 2891 CCATGTTATTCCTTTCAGCTTGCAATCAGAAATACTTATAGCCGACAGCAANTTCATTAGAA 2950
QY 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
DB 2951 AAAAAAAGGCTTCCCTTCTGCAGACCCAGGCTCAGTGTCTGAGAGGGCGGTATGTGTCTC 3010
QY 1001 AsnCysProLeuGlyThrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
DB 3011 AATTGCCCTTTGGGACCTTATTAATCTGGAACATTTTCACTGTGAAAGCTGCCGGATC 3070
QY 1021 GlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
DB 3071 GGATCTTATCAAGTGAAGAGGGGCACTTGAGTGCAGAGCTTTGCCCTCTGGGATGTAC 3130
QY 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
DB 3131 ACGGAATATATCCATTCAGAAAAATCTCTGATTTGTAAGCTCAGTGTAAACAAGGACCC 3190
QY 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
DB 3191 TACTCATACAGTGGACTTGAGACTTGAATCGGTCCACTGGGCACCTTATCAGCCAAA 3250
QY 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
DB 3251 TTTGGTTCGCGAGCTGCCCTCTGCTGTCCAGAAAAACACTCAACTGTGNAAGAGGAGGCC 3310
QY 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
DB 3311 GTGAACATTTCTGATGTGGAGTTCTTGTTCACAGAGAAAAATTTCTCGCGTCTCTGGGTTA 3370
QY 1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
DB 3371 ATGCCCTGTCCACCATGTCTCTCGTACTATTTACCACTAATGAGGAGGAGCCCTTCTGCG 3430
QY 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160

Db 3431 |||||CTGGCCCTGTCCTTTATGGAACTACCCCATTCCTGGTCCAGATCCATCACGAATGT||| 3490
Qy 1161 SerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu||| 1180
Db 3491 TCAAGTTTATGTTCACTTCTCAGCGGACAGGAAAGTGTGGTCCCTGCTCTCTT 3550
Qy 1181 GlyHisIleTyrArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
Db 3551 GGACATATTAAGAGAGCATGAATCAGCAGTCAGGTTTCCATGAATGCTTCTTTAAAC 3610
Qy 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
Db 3611 CCTTGGCCCAATAGTGGAACTCGCCAGCACTTGGCGTGTGTATGTTGTCTCTGTCCA 3670
Qy 1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspGluCysSerProLeuProCys 1240
Db 3671 CTTCGATATACAGGCTTAAGTGTGAACAGACATCGATGAGTGCAGCCCATGCTTTC 3730
Qy 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
Db 3731 CTCACCAATGAGTGTGAAGACCTAGTGTGGGAATTCATTGTGAGTGCCTTCAGGT 3790
Qy 1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsn 1280
Db 3791 TACACAGGTTCAGCGGTGTGAAGAAATATAATGATGATGATGCTCCAGTCTTGTTAAT 3850
Qy 1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
Db 3851 AAAGGAATCTGTGTGATGTGTGGCTGCTATCGTTCACATGTTGTAAGAGATTGTA 3910
Qy 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
Db 3911 GGCCTGCATTCGAAACAGAAAGTCAATGAATGCCAGTCAAAACCCATGCTTAATATGCA 3970
Qy 1321 ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
Db 3971 GTCTGTGAAGACCAAGTGTGGGGATCTTGTGCAAAATGCCACCTCGGATTTTGGGTACC 4030
Qy 1341 ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
Db 4031 CGATGTGGAAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGCTACCTGT 4090
Qy 1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db 4091 AAAGACGGTGGCAATAGCTTCAGATGCCCTGTGTGCGAGCTGGCTTCACAGGATCACACTGT 4150
Qy 1381 GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db 4151 GAATGAACATCATGAATGTGAGTCTAATTCATGTAGAAATCAGGCCACCTGTGTGGAT 4210
Qy 1401 GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db 4211 GAATTAATTCATACAGTTGTAATGTTCAGCCAGGATTTTCAGGCAAAAGGTGTGAACA 4270
Qy 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db 4271 GAACAGCTTCACAGGCTTTAAACCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTGATG 4330
Qy 1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSerSer 1460
Db 4331 CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACTGTACCTTCTGAGTGAATTCCTCT 4390
Qy 1461 AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db 4391 GACGACATGAACACTTGAACACCAATCTCTATGCGATGTTGATAACGGCAGCACAATACC 4450
Qy 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db 4451 TTGCTCTGACTGATTAATACGGCTGGGTCTCTTTATGTGAATGGCAGGGAAGATAACA 4510
Qy 1501 AsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrThrSerAla 1520
|||

RESULT 2

US-10-150-821-1

; Sequence 1, Application US/10150821

Db 4511 AACCTGTCCTCGGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGACAAAGTGCC 4570
Qy 1521 AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db 4571 AATGCCATCTGGAAAGTCTATATCGATCGGAAAATATCTGACGGTGGTGTGGCCTCTCT 4630
Qy 1541 ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluGluAspLysLys 1560
Db 4631 GTTGGTTTGGCCCATACCTCGGTGGTGGTGTAGTTCCTGGGCAAGAGCAAGCAAAA 4690
Qy 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
Db 4691 GGAGAGGATTCAGCCAGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG 4750
Qy 1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db 4751 GACTATGTCTGTCTCCACAGCAGTGAAGTCACTGGCTACCTCTGCCAGAGAACTC 4810
Qy 1601 SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db 4811 AGTAAAGAAACGTGTTAGCATGGCTGATTTCTTGTCAAGAAATTTGGGGAAAGTGAAG 4870
Qy 1621 IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 4871 ATCGATTCTAAGAGCATATTTTGTCTGATTGCCACGCTTAGGAGGGTCACTGCCTCAT 4930
Qy 1641 LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 4931 CTGAGAACTGCATCTCGAAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTCTCTGTGATCCA 4990
Qy 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
Db 4991 GCTTCCAGCTGTGTCGGAAACCTGTGCAGTACTGTCTGAATCAAGGACAGTGGACACAA 5050
Qy 1681 ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db 5051 CCACCTCTCTACTGTGAAGCATTTAGTGTGGGTGCCACCTCTCTTGAGAAATGGCTTC 5110
Qy 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db 5111 CATTCAGCGATGACTTCTATGCTGGCAGCACAGTAACCTACACGTGCACCAATGGCTAC 5170
Qy 1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSer 1740
Db 5171 TATCTATTGTGGTGAATCAAGGATGTTCTGTACAGATATATGGAGCTGGGAACGGGTTC 5230
Qy 1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5231 CCATCTCTGCTTGTATGTCGATGAGTGTGCAGTTGGATCAGATTGTAGTGAGCATGCTTCT 5290
Qy 1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db 5291 TGCCCTGAACGTAGATGGATCTCTACATATGTTCTGTCCTCCCGTACACAGGATGGG 5350
Qy 1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5351 AAAAACTGTCAGAACCTTATAAATATGAAGGCTTCAGGAAATCCGGAATATGGCCACTCC 5410
Qy 1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db 5411 TCAGGTGAGATTTATACAGTAGGTGCGGAAGTCAATTTTGTGTCTCAGGAAGGATACCAG 5470
Qy 1821 LeuMetGlyValThrIleThrCysLeuGluSerGlyGluThrAsnHisLeuIlePro 1840
Db 5471 TTGATGGGAGTAAACAAAATCAATGTTTGGAGTCTGGAGAAATGGAAATCATCTAATACCA 5530
Qy 1841 TyrCys 1842
Db 5531 TATTGT 5536

Qy 581 GlnAspSerAlaAenValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
 Db 1751 CAAGATTCTGCAATGTTACCTGGCAGATTCCAAACAGCTAAAGACAACCTCTGGTGAAG 1810
 Qy 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
 Db 1811 GTGTGAGTCCAGTTCATCAGCTTTACCCACCTTACCTTTTCCAGTGGAGATGTT 1870
 Qy 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
 Db 1871 GCTATGTAATACACGGCACTGACCTATCCGCAACACAGGCCAGCTGCAATTTCCATATC 1930
 Qy 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
 Db 1931 AAGGTATTGATGCAGAACCACTGTCATAGACTGGTCAGACTCTCCACCTCCGCTCCAG 1990
 Qy 661 ValSerGluLysValHisAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
 Db 1991 GTCTGGGAGAGGTACATGCCGCAAGCTGGGATGAGCTCAGTTCTCAGACAACTCAGG 2050
 Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
 Db 2051 GCTGAATTGGTCATTTACCAAGATCATACACAGGAGACCTTTCCCTCAAGGGGAGACT 2110
 Qy 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
 Db 2111 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATT 2170
 Qy 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
 Db 2171 GTCATAAAGGTTCCTCCCTGTGAATTCATTCACACCTGTAAATGGGATTTTATATGC 2230
 Qy 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
 Db 2231 ACTCCAGATATACCTGGAGTCACTGTACATTAACCTTGGAGGGCTATGATTTTCA 2290
 Qy 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
 Db 2291 GAAGGGTCTACTGACAAATATTATTGTCTTATGAAGATGGCGTCTGGAAACCAACATAT 2350
 Qy 781 ThrThrGluTrpProAspCysAlaLysLysValArgPheAlaAsnHisGlyPheLysSerPhe 800
 Db 2351 ACCACTGAATGGCCAGACTGTGCCCCAAAACGTTTGTCTAACCCAGGGTTCAAGTCTTT 2410
 Qy 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
 Db 2411 GAGATGTTCTACNAAGCAGCTGTGTGATGACACAGATCTGATCAAGAAAGTTTTCTGAA 2470
 Qy 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
 Db 2471 GCATTTGAGACGACCTGGGAAAAATGGTCCCATCATTTTGTAGTGATGCGAGGACATTT 2530
 Qy 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
 Db 2531 GACTGCAGACTGGAGAGACCTGACCAAAATATTTGCTAGATATATTAATATGACTAT 2590
 Qy 861 GluAsnGlyPheAlaIleGlyProGlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSer 880
 Db 2591 GAAAAATGGCTTTGCAATTTGGACCACTGGCTGGGGTGGCAGCTAAATAGGCTGGATTA 2650
 Qy 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
 Db 2651 TACGATGATCTTCTGACACTGTGCAGAAACAGCCCAAGCAATCGCAATGCCAAGTCC 2710
 Qy 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
 Db 2711 TCACGGATTAAGAAAGTGGCCCATATCTGACTATTAATAATTAAGTTAATTTTAAACATC 2770
 Qy 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGln 940
 Db 2771 ACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAA 2830

Qy 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
 Db 2831 CGACTCTTCAGACATTCGAAACATATCAAAATAAAAGGAAAGGACTCTCAACAAAGAC 2890
 Qy 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGlu 980
 Db 2891 CCCATGATATTCCTTCAGCTTGCATCAGAAATATTTATAGCCGACAGCAATTCATTAGAA 2950
 Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
 Db 2951 ACMAAAAGGGCTTCCCTCTTCGACAGCAGGCTCAGTGTCTGAGAGGGGTATGTGTCTC 3010
 Qy 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
 Db 3011 AATTGGCCCTTTTGGGAACCTATTATTAATCTGGAACATTTTCACTGTGAAAGCTGCCGGATC 3070
 Qy 1021 GlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
 Db 3071 GGATCTATCAAGATGAAGAGGGCACTTGATGCAAGCTTTGCCCTCTGGGATGTAC 3130
 Qy 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
 Db 3131 ACAGAAATATATCCATTCAGAAACATCTCTGATGTAAAGCTCAGTGTAAACAGAGCAC 3190
 Qy 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
 Db 3191 TACTCATAGTGGACTTGGAGCTTGTGATTCGTCTCACTGGGCACTTATCAGCCAAA 3250
 Qy 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
 Db 3251 TTTGGTTCGCCGAGCTGCTCTCGTGTCCAGAAAACACCTCACTGTGAAAAGAGAGGCC 3310
 Qy 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
 Db 3311 GTGAACATTTCTGCATGTGGAGTTCCTTGTCCAGAAAGGAAATTTCTCGGGTTCTGGGGTA 3370
 Qy 1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
 Db 3371 ATGCCCTGTACCCCATGCTCTGTGACTATTAACCACTTAATGACGGGAAGGCTTCTGC 3430
 Qy 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
 Db 3431 CTGGCTGTCTCCCTTTATGGAATACCTCCCATTTGGTGTCCAGATCCATCAGAAATGT 3490
 Qy 1161 SerSerPheSerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu 1180
 Db 3491 TCAAGTTTTAGTTCAACTTTCTCAGCGCAGAGAAAGTGTGGTGGCCCTCTCTCTT 3550
 Qy 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
 Db 3551 GGACATATTTAAAGAGGCGATGAATCAGCACTCAGTGTTCATGATGCTTCTTTAAC 3610
 Qy 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
 Db 3611 CTTGGCCCAATAGTAGGAACCTGCCAGCACTTGGCGGTGTGTATGTTGTCTCTGTCCA 3670
 Qy 1221 LeuGlyTyrThrGlyLysLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
 Db 3671 CTTGGATATACAGGCTTAAAGTGTGAACAGACATCATGATGAGTGCAGCCACTTCTGC 3730
 Qy 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
 Db 3731 CTCACATATGGGTTGTGAAGACCTAGTTGGGAAATTCATTTGTGAGTGGCCATCAGGT 3790
 Qy 1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
 Db 3791 TACACAGTTCAGCGGTGTGAAGAAATATAAATAGATGTAGTGTCCAGTCTCTGTTTAAAT 3850
 Qy 1281 LysGlyIleCysValAspGlyValalaglyTyrArgCysThrCysValLysGlyPheVal 1300
 Db 3851 AAAGGAATCTGTGTGATGGTGTGGCTGTGCTATGTTGACATGTGTGAAGAGATTTGTA 3910
 Qy 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320

Db 3911 GGCCTGCATTTGTGAACAGAACTCAATGAATCCAGTCAAAACCCATGCTTAAATAATGCA 3970
Qy ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
Db 3971 GTCTGTGAAGACAGGTTGGGGGATTTCTTGTGCAATGCCCACCTGGATTTTGGGTACC 4030
Qy ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
Db 4031 CGATGTGGAAAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGAGACTACTGT 4090
Qy LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db 4091 AAAGACGGTCCCAATAGCTTCAGATGCTGTGTGACAGTGGCTTCACAGGATCACACTGT 4150
Qy GluLeuAsnLysGlnCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db 4151 GAAATTGAACATCAATCAATGTCAGTCTAATCCATGATAGAAATCAGGCCACCTGTGTGGAT 4210
Qy GluLeuAsnSerThrCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db 4211 GAAATTAAATTTCAACAGTTGTAAATGTCAGCCAGGATTTTCAGGCAAAAGGTGTGAACA 4270
Qy GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db 4271 GAACAGTCTACAGGCTTTAACTGGATTTTGAAGTTCTGGCATCTATGGATATGTCATG 4330
Qy LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer 1460
Db 4331 CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACTGCTACCTTCTGGATGAATCTCT 4390
Qy AspAspMetAsnTyrGlyThrProLysSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db 4391 GACGACATGACATGGAACACCAATCTCTATGTCAGTTGATTAACGCGCAGCAATACC 4450
Qy LeuLeuLeuThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db 4451 TTGCTCTGACTGATTTAAACGCTGGTTCCTTATGTCAGTGGCAAGCAAGATAACA 4510
Qy AsnCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAla 1520
Db 4511 AACTGTCCCTCGGTGAATGATGCGAGATGTCATATATGCAATCACTTGGACAAGTGC 4570
Qy AsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db 4571 AATGGCATCTGGAAGTCTATATCGATGGGAATTTATCTGACGGTGTGCTGCCCTCTCT 4630
Qy ValGlyLeuProLysProGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLys 1560
Db 4631 GTTGGTTTCCCATACCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4690
Qy GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp 1580
Db 4691 GGAGAGGGATTCAGCCAGCTGAGTCTTTTGTGGCTCCATAAGCCAGCTCAACCTCTGG 4750
Qy AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db 4751 GACTATGTCTGTCTCCACAGCAGGTGAAGTCACTGCTTACCTCTGCTGCCAGAGGAATC 4810
Qy SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db 4811 AGTAAAGAAACAGTGTAGCATGGCTGATTTCTTGTGAGGAATGTGGGAAGTGAAG 4870
Qy IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 4871 ATCGATTCGAAGACATATTTTGTCTGATGTCAGCTTCCAGCTTAGGAGGTGAGTGCCTCAT 4930
Qy LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 4931 CTGAGAATCTGCACTGAAGATTTTAAAGCCAGGTTCCAAGTCAATCTGTTCTGTATCCA 4990
Qy GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680

Db 4991 GGCTCCAGCTGCTCGGGAACCCCTGTGCAGTACTGTCTGAATCAAGACAGTGGACACAA 5050
Qy ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db 5051 CCATCTCTCCATCTGTAAGACCATTAAGTGTGGGGTGCCACTCTCTTTGGAGAAATGGCTTC 5110
Qy HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db 5111 CATTCAGCCGATGACTTCTATGCTGGCAGCACAGTAACCTACCACTGTCACCAATGGCTAC 5170
Qy TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer 1740
Db 5171 TATCTATGGGTGACTCAAGGATGTTCTGTACAGATAATGGAGCTGGAAACGGCTTTCA 5230
Qy ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5231 CCATCTGCTTGTGATGTCAGTGTGTCAGTGTGATGATGATGATGATGATGATGATGAT 5290
Qy CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db 5291 TGCTGACGCTAGATGATCTCATATATGTTCAATGTGTCCACCGTACACAGAGATGGG 5350
Qy LysAsnCysAlaGluProLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5351 AAAAATCTGTCAGAACCTATAAATGTAAAGCTCCAGGAAATCCGAAATGGCCACTCC 5410
Qy SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGlyTyrGln 1820
Db 5411 TCAGGTGAGATTTATACAGTAGGTGCGCAAGTGCATTTTGTGTGAGGAAGTATCCAG 5470
Qy LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 5471 TTGATGGAGTACCAAAATCACAATGTTTGGAGTCTGAGAAATGGAATCATCTAATACCA 5530
Qy TyrCys 1842
Db 5531 TATTTGT 5536

RESULT 3
US-10-295-027-429
; Sequence 429, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevesi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349

; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 429
 ; LENGTH: 10674
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-429

Alignment Scores:

Pred. No.: 0 Length: 10674
 Score: 10047.00 Matches: 1812
 Percent Similarity: 98.86% Conservative: 9
 Best Local Similarity: 98.37% Mismatches: 13
 Query Match: 98.32% Indels: 8
 DB: 16 Gaps: 2

US-09-977-053-6 (1-1842) x US-10-295-027-429 (1-10674)

Qy 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
 Db 1 ATGTGGCCCTCGCCCTGGCCCTTTGTTGGCTGGGGTCTGGGGCTGCTTTTCGGGGCTGGGGGACC 60
 Qy 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
 Db 61 TTTTCAGAGAGTGTCCCGTGGGCGAATTCAGTTCCGGCTCTTCCCGGAGACCCGGGCC 120
 Qy 41 GlyAlaProGlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArg 60
 Db 121 GGGGCCCGCGGAGTATCCCGCGCGCGCGCTCTCGCGAGCGAGCGCGCGGAGCAGA 180
 Qy 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
 Db 181 GTGGAGCGGCTGGGCGAGCGTTCGGCGAGCGTGGCGGCTGCTCGGGGAGCTCAGCGAG 240
 Qy 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg 100
 Db 241 CGCTGGAGCTTGCTTCTGCTGTGATGATTCGTCCAGCGTGGGCGAAGTCAACTTCGCG 300
 Qy 101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
 Db 301 AGCGAGCTCATGTTCTGTCGCAAGCTGCTGTCGACATCCCGTGGTGGTCCCGCGCCAGC 360
 Qy 121 ArgValAlaIleValThrPheSerSerLysSerValValProArgValAspTyrIle 140
 Db 361 CGCTGGCCATCGTGACCTTCTGTCGAAGAACTACGTGGTGGCGCGGTGATTTACATC 420
 Qy 141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGlnIleProAlaIle 160
 Db 421 TCCACCGCGCGCGCGCAGCAGTGGCGCTGCTCTCCAGAGATCCCTGCCATC 480
 Qy 161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeu 180
 Db 481 TCCTACCGAGGTGGCGGACCTTACCAAGGCGGCTTCCAGCAAGCGCGCAAAATTCCT 540
 Qy 181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuLeuThrAspGlyTyrSerAsn 200
 Db 541 CTTTCATGCTAGAGAAATCTCAACAAAGTTGTATTTCTCATCATCTGATGGATTTCCAAAT 600
 Qy 201 GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
 Db 601 GGGGAGACCTTAGACCAATTCACGGTCACTGCGAGATTCAGGAGTGGAGATCTTCACT 660
 Qy 221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
 Db 661 TTTGGCATATGCAAGGAAACATTCCAGAGCTGGAATGATGCTTCCACCCCAAGGAG 720
 Qy 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260

Db 721 GAGCACTGTTACCTGCTACACAGTTTGGAGAAATTTAGGCTTTAGCTGCCGGGATTC 780
 Qy 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
 Db 781 CATGAAGATCTACCTTCTGGGAGTTTATTTCAAGATGATATGTCCTCACTGCTCATATCTT 840
 Qy 281 CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
 Db 841 TGTGATGAAGCAAGGAGCTGCTGTGACCGAATGGGAAGTGCANATGTGGGACACACACA 900
 Qy 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
 Db 901 GGCCATTTTGAGTGCATCTGTGAAAGGGGTATTACGGGAAGGCTCTGCAGTATGATATGC 960
 Qy 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
 Db 961 ACAGCTCTCCCATCGGCGACATACAAACCTGAAGGCTCACACGAGGAGAAATCAGCAGTGC 1020
 Qy 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
 Db 1021 ATTCCATGTCCTGATGAAATCACACCTCTCCACCTGGGAAGCACATCCCTCTGAAGACTGT 1080
 Qy 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
 Db 1081 GTCTCAGAGAGGATACAGGCACTGTGCCAGAGCTGTGAACCTGTCCACTGCGCTGCC 1140
 Qy 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
 Db 1141 CTGAAGCCTCCCGAAATGGTTACTTTATTCAAACACCTTGCACCAACCACTTCAATGCA 1200
 Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
 Db 1201 GCCTGTGGGGTCCGATGTCACTGGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1260
 Qy 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
 Db 1261 CTACCAATGTTTGTGTGCTCGGTTCCAGAGACTACTGCAGAGTAAGAACAATGTTATATA 1320
 Qy 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
 Db 1321 CTCCGCCAGCGAACAATGCGCCATCAGCTGTTCTTACAAGGGAAATGTTATATAAGACA 1380
 Qy 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
 Db 1381 ACATGTTTGGTTGCTCTGTGATGAGGTTACAGACTAGAGGCGAGTGTAAAGCTTACTTGT 1440
 Qy 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
 Db 1441 CAAGGAACAGCCAGTGGGATGGCGCAGAACCCCGGTGTGTGGAGCGCCACTGTGTCACC 1500
 Qy 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
 Db 1501 TTTTCAGATGCCCAAGATGTCATCATATATATATATATATATATATATATATATATATAT 1560
 Qy 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
 Db 1561 TTTGGAGCAATCTGCTATGTAAGTTGTCGCCCAAGGGTTCAATTTATCTGGAGTCAAGAA 1620
 Qy 541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
 Db 1621 ATGCTGAGATGTACCACTTCTGGAATAATGGAATGTCGAGTTCAGGAGCTGTGTGATAA 1680
 Qy 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
 Db 1681 GACGTGGAGGCTCTCAAAATCAACTGTCTTAGGACATAGAGGCTTAAGACTCTGGAAACAG 1740
 Qy 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
 Db 1741 CAAGATTCTGCCAATGTTACCTGGCAGATTCACAGAGCTTAAGACAACTCTGTGTGAAG 1800
 Qy 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620

1801 GTGTCACTCCAGCTTCATCCAGCTTTTCCACCCACCTTACCTTTTCCCAATTCAGATGTT 1860
Db
621 AlalleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
Qy
1861 GCTATCGTATACAGCGCAACTGACCTTATCCGCAACCGCCAGCTGCATTTTCCATATC 1920
Db
641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
Qy
1921 AGGTTATTGATGCAGACCACTGTTCATAGACTGGTGAGATCTCCACCTCCGCTCCAG 1980
Db
661 ValSerGluLysValHisAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
Qy
1981 GTCTCGAGAGGTACATGCGCAAGCTGGGATCAGCTCAGTTCTCAGACAACTCAGGG 2040
Db
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
Qy
2041 GCTGAATTTGGTCAATACCAAGTCTATACCAAGGAGACCTTTTCCCTCAAGGGGAGACT 2100
Db
701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Qy
2101 ATAGTACAGTATACAGCCACTGACCTCAGGCAATAAACAGGACATGTGATATCCATAT 2160
Db
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
Qy
2161 GTCATAAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGGATTTTATATGC 2220
Db
741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
Qy
2221 ACTCCAGATAAATCTGGAGTCAACTGTACATTAATCTTGTGGAGGGCTATGATTTTCA 2280
Db
761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
Qy
2281 GAAGGCTCTACTGACAGATATATTTGTGCTTATGAGATGGCGTCTGGERAACCAATAT 2340
Db
781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Qy
2341 ACCACTGAATGGCAGACTGTGCAAAAACGTTTTCGCAAAACCAACGAGTCTTCAAGTCTTT 2400
Db
801 GluMetPheTyrIlysalAlaArgCysAspAspThrAspLeuMetIlyLysPheSerGlu 820
Qy
2401 GAGATGTTCTTACAAAGCAGCTCGTTGTGATGCACAGATCTGATGAAGAGATTTTCTGAA 2460
Db
821 AlaPheGluThrThrLeuGluLysMetValProSerPheCysSerAspAlaGluAspIle 840
Qy
2461 GCATTTGACAGCCCTCGGAAAAATGGTCCCATCATTTGTAGTGTATGACAGGACATT 2520
Db
841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
Qy
2521 GACTGCAGACTGGAGGAGAACCTGACCAAAAAATATTGCTAGAAATATAATTTATGACTAT 2580
Db
861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAlaAsnArgLeuAspTyrSer 880
Qy
2581 GAAATGGCTTTTGCAATTTGACACAGTGGCTGGGGTGCAGCTAATAGGCTGGATTAATCT 2640
Db
881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
Qy
2641 TAGCATGACTTCTTGACACTGTGCAAGAAACAGCCACAGCAATCGGCATGCAAGTCC 2700
Db
901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
Qy
2701 TCACGGATTTAAAGAGTGCCTTATCTGACTATAAAATTAAGTTAAATTTTAAACATC 2760
Db
921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln 940
Qy
2761 ACAGCTAGTGTGCCATTTCGCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAA 2820
Db
941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
Qy
2821 CGACTCTTCAGACATGGAACTATACAAATAACTGAAAGGACTCTCAACAAGAC 2880
Db
961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
Qy
2881 CCCATGTAATTCCTTTTTCAGCTTGCATCAGAAATACTTATAGCCGACAGCAATTCATTAGAA 2940
Db

Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
Db 2941 ACAAAAAGAGCTTTCCCTTCTGCAGACAGGCTCAGTGTCTGAGAGGGCGTATGTGTCTC 3000
Qy 1001 AspCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
Db 3001 AATTGGCCCTTTTGGAACTTATTAATCTGGAACATTTTCACTGTGAAAGCTGCGGATC 3060
Qy 1021 GlySerTyrClnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
Db 3061 GGAATCTTATCAGATGAAGAGGCAACTTGAGTGCAGGCTTTGCCCTCTGGAATGTAC 3120
Qy 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
Db 3121 ACAGATATATCATTCATCAGAAACATCTCTGATTTGTAAAGCTCAGTGTAAACAGGAC 3180
Qy 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGluProLys 1080
Db 3181 TACTCATACAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCACTTATCAGCCAAA 3240
Qy 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
Db 3241 TTTGGTTCCCGAGCTGCTCTCGTGTCCAGAAAACACCTCAACTGTGAAAAGAGGACC 3300
Qy 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
Db 3301 GTGAACATTTCTGCATGTGGAGTTCTTGTTCAGAGAGAAATTTCTCGGTTCGGGTTA 3360
Qy 1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
Db 3361 ATGCCCTGTCAACCATGTCTCTGTGACTATTACCAACTAATGAGGAGAGGCTTCTGCG 3420
Qy 1141 LeuAlaCysProPheTyrGlyThrThrProPhealaglySerArgSerIleThrGluCys 1160
Db 3421 CTGCGCTGTGCTCTTTTATGGAACCTACCCCATTCGCTGGTTCCAGATCCATCAGCAATGT 3480
Qy 1161 SerSerPheSerSerThrPheSerAlaAlaGluSerValValProProAlaSerLeu 1180
Db 3481 TCACCT-----TCAGTCTGAAATATTAATTTTTCGGTGGATTT 3519
Qy 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
Db 3520 GGGCATCTGGAGTTGTTAAAT---TGTCTCTCTGAGGTTTTTCCATGAATGCTTCTTTAAC 3576
Qy 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
Db 3577 CCTTGGCACAATAGTGGAACTTCCAGCAACTTGGGGCTGGTTATGTTGTCTCTGTCCA 3636
Qy 1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
Db 3637 CTGGAATATACAGCTTAAAGTGTGAACAGACATCGATGAGTGCAGCCCACTGCTTGC 3696
Qy 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyLysPheIleCysGluCysProSerGly 1260
Db 3697 CTCAACAAATGGAGTTTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGGCCATCAGT 3756
Qy 1261 TyrThrGlyClnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
Db 3757 TACACAGGTGAGCGGTGTGAAGAAATATAAATGAGTGTAGCTCCAGTCTCTGTTAAAT 3816
Qy 1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
Db 3817 AAGGAATCTGTGTTGATGTTGGCTGGCTATGCTGTGCATGTGTGAAGAGATTTGTA 3876
Qy 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
Db 3877 GGCCTGCATTTGTGAACAGAGAGTCAATGAATGCAGTCAAAACCCATGCTTAAATTAATGCA 3936
Qy 1321 ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
Db 3937 GTCTGTGAAGACAGGTTGGGGATTTCTTGTGTCAAAATGCCACCTGGAATTTTGGGTACC 3996

```
Qy 1341 ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValThrCys 1360
Db 3997 CGATGTGGAAAGACGTCGATGTCCTCAGTCAGCCATGCAAAAATGGAGCTACCTCTGT 4056
Qy 1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db 4057 AAGACGGTGCATAGCTTCAGATGCTGTGTGTCAGCTGCTTCACAGGATCACACTGT 4116
Qy 1381 GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db 4117 GAATTGAACATCAATGAATGTCAGTCTAATCATCTAGAAATCAGGCCACCTGTGTGGAT 4176
Qy 1401 GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysLysThr 1420
Db 4177 GAATTAAATTCATACAGTTGTAATGTCGCGCAGGATTTTCAGGCAGAAAGGTGTGAACA 4236
Qy 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db 4237 GAACAGTCTACAGGCTTTAACTGGAATTTTGAAGTTTCTGGCATCTATGGATATGTCATG 4296
Qy 1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer 1460
Db 4297 CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACTGTACCTTTCTGGATGAATCTCTCT 4356
Qy 1461 AspAspMetAsnTyrGlyThrProLysSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db 4357 GACGACATGAATCTATGGAACACCAATCTCTATGCGGTGTGATTAACGGCAGCGACATACC 4416
Qy 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db 4417 TTGCTCCTGACTGATTAACCGCTGGGTCTTTATGTGAATGGCAGGAAAGATAACA 4476
Qy 1501 AsnCysProSerValAsnAspGlyArgTrpHisIleIleAlaIleThrTrpThrSerAla 1520
Db 4477 AACTGTCTCCCTCGTGAATGATGGCAGATGGCATCATATTTGCATTCACCTTGGACAAAGTCC 4536
Qy 1521 AsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyValGlyLeuSer 1540
Db 4537 AATGGCATCTGGAAGTCTATATCATGGGAATATCTGACGGTGTGCTGGCTCTCTCT 4596
Qy 1541 ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLys 1560
Db 4597 GTTGGTTTGGCCATACCTGGTGTGTGGTGTAGTTCTCTGGGGCAGAGCAAGACAAAAA 4656
Qy 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp 1580
Db 4657 GGAGAGGATTCAGCCCACTGAGTCTTTTGTGGCTCCATGAAGCCAGCTCAACCTCTGG 4716
Qy 1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db 4717 GACTATGTCCTGCTCCACAGAGTGAAGTCACTGGCTACTCTCTGCCAGAGGAACCTC 4776
Qy 1601 SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db 4777 AGTAAAGGAAACGTTGTAGCATGGCTGATTTCTTGTTCAGGAATTTGTGGGAAAGTGAAG 4836
Qy 1621 IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 4837 ATCGATTCTAAGAGCATATTTGTCTGATTTGCCACCGCTTAGGAGGGTCACTGCTCTCAT 4896
Qy 1641 LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 4897 CTGAGACTGATCTGAAGATTTAAAGCCAGTTCCAAAGTCAATCTGCTCTGTGATCCA 4956
Qy 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln 1680
Db 4957 GGCCTCCAGCTGGTGGGAAACCTGTGTCAGTACTGCTCTGAATCAAGGACAGTGGACACA 5016
Qy 1681 ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db 5017 CCACCTTCCTCATCTGTGAACGATAGCTGTGGGGTGGCCACTCTCTTTCAGGAATGGCTTC 5076
Qy 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720

|||||
5077 CATTCAGCCGATGATCTTCTATGCTGGCAGCAGTAACCTACCACTGCAACCAATGGCTAC 5136
Qy 1721 TyrIleuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer 1740
Db 5137 TATCTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCGTTTCA 5196
Qy 1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5197 CCATCTGCTTGTATGTCATGATGATGTCAGTTCAGTTCAGATTGTAGTCAGCATGCTTCT 5256
Qy 1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrTrpThrGlyAspGly 1780
Db 5257 TGCCTGACGTTAGATGATCTTACATATGTTTCATGTCTCCACCGTACACAGGAGATGGG 5316
Qy 1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5317 AAAAATCTGTGAGAACCTTATAAATGTAAGGCTCCAGAAATCCGGAAATGGCCACTCC 5376
Qy 1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db 5377 TCAGTGTGATTTATACAGTAGTGGCGAGTCACTTTTGTGTCCAGGAAGATACCAAG 5436
Qy 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 5437 TTGATGGGAGTAACCAAAATCACATGTTTGGAGTCTCGAGAATGGAATCATCTAATACCA 5496
Qy 1841 TyrCys 1842
Db 5497 TATTGT 5502

RESULT 4
US-10-295-027-1080
; Sequence 1080, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
```


1921 AAGGTTATTGATGAGAACCACTGTCATAGACTGGTGCAGATCTCCACCTCCCGTCCAG 1980
661 ValSerGluLysValHisAlaAAserTrpAspGluProGlnPheSerAspAsnSerGly 680
1981 GTCTCGGAGAGGTACATGCCGCAAGCTGGGATGAGCTCAGTCTCTCAGACAACTCAGG 2040
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
2041 GCTGAATTGGTCATTACCAAGAGTCATACACAGGAGACCTTTTCCCTCAAGGGGAGACT 2100
701 IleValGlnThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
2101 ATAGTACAGTATACAGCACCTGACCCCTCAGCGCAATACAGACACATGTGATATCCATAT 2160
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
2161 GTCATAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGGATTTATATGC 2220
741 ThrProAspAsnThrGluValAsnCysThrLeuThrCysLeuGluGlyThrAspPheThr 760
2221 ACTCCAGATAATCTGGAGTCACTGTATCACTTAACCTTGGAGGGCTATGATTTTCA 2280
761 GluGlySerThrAspLysTyThrCysAlaTyGluAspGlyValTyrLysProThrTyr 780
2281 GAAGGCTCTACTGACAGTATATTTGTCTTATGAGATGGGCTCTGGAACCAACATAT 2340
781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
2341 ACCACTGAATGCCAGACTGTGCCAAAAAAGCTTTTGCAACCAACGCGGTCAAGTCCCTTT 2400
801 GluMetPheTyLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
2401 GAGATGTTCTCAAGAGCAGCTGTTGTGTATGACACAGATCTGATGAAGAAGTTTCTGAA 2460
821 AlaPheGluThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
2461 GCATTTGAGACACCTCGGAAATATGTCCCATCATTTTGTAGTATGACAGGACAT 2520
841 AspCysArgLeuGluGluAsnLeuThrLysLysTyCysLeuGluTyAsnTyrAspTyr 860
2521 GACTGCAGACTGGAGGAACTGTACCAAAAAATATGCTTAGAATAATAATATGACTAT 2580
861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
2581 GAAATATGCTTTGCAATTTGGAACCAAGTGGCTGGGTCAGCTAAATAGGCTGATTAATCT 2640
881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
2641 TACGATGACTTCTTGACACTGTGCAAGAAACAGCCACAGCATCGGCCAATGCCAAGTCC 2700
901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
2701 TCAGGATTAAGAAGTGGCCCATTAATCTGACTATAAAATTAAGTTAAATTTTAACATC 2760
921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
2761 ACAGCTAGTGTGCCATTAACCGATGAAGAATGATACCTTTGAATGGGAATCAGCAA 2820
941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
2821 CGACTCTCTCAGACATTTGGAACTATCACAATAAATCACTGAAAGGACTCTCAACAAAGAC 2880
961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
2881 CCCATGATATCTTTCAGCTTGCATCAGAAATATCTATAGCCGACAGCAATTCATTAGAA 2940
981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
2941 ACAAAAAGGCTTCCCTCTTCGACACAGGCTCAGTCTGAGAGGGCGTATGTGTCT 3000
1001 AsnCysProLeuGlyThrTyrTrpAsnLeuGluHisPheThrCysGluSerCysArgIle 1020

3001 AATTGCCCTTTGGGAACCTATTATAATCTGCAACATTTTCACTGTGAAAAGCTGCCGATC 3060
1021 GlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
3061 GGATCTCTATCAAGATGAAGAAGGGCACTTGAGTGCAAGCTTTGCCCTCTCGGAGTAC 3120
1041 ThrGluTyIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
3121 ACGGATATATCCATTCCAGAAACATCTCTGATTTGAAGCTCACTGTAAACAAGGCACC 3180
1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
3181 TACTCATACAGTGCAGCTTGAGACTTGTGATCTGTCCTCACTGGGCATTTATCAGCCAAA 3240
1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
3241 TTTGGTTCCCGAGCTGCTCTCTGTGTCAGAAAACCTCAACTGTGAAAAGAGAGGCC 3300
1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
3301 GTGAACATTTCTGCATGTGGAGTTCCTTGTCCAGAAGGAAATTTCTCGGTTCTGGTTA 3360
1121 MetProCysHisProCysProAspAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
3361 ATGCCCTGTCCACCATGCTCTGCTGACTATTACCAACCTTAATGSCAGGGAAGCCCTTCTGC 3420
1141 LeuAlaCysProPheTyGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
3421 CTGGCTGTCTCTTTATGGAACTACCCCATTCGCTGGTTCAGATCCATCAGAAATGT 3480
1161 SerSerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeu 1180
3481 TCNACT-----TCAGTTCTGAATATTACTATTTTTCGTGGATTT 3519
1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
3520 GGGCATCTGGAGTTGTATAAT---TGTCTCTCTGAGGTTTTCATGAATGCTTCTTAAC 3576
1201 ProCysHisLeuSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
3577 CTTGCCCAATAGTAGGAACCTGCCAGCACTTGGCGGTGGTTATGTTTGTCTGTCCA 3636
1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
3637 CTGGATATACAGGCTTAAAGTGTGAACAGACATCGATGAGTGCAGCCCTGCTTGC 3696
1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
3697 CTCACAAATGGAGTTGTAAAGACCCTAGTTGGGGAATTCATTTGTGAGTGCCTCAGGT 3756
1261 TyrThrGlyGlnArgCysGluLeuAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
3757 TACACAGTTCAGCGGTGTGAAGAAATATAAATGAGTGTAGCTCCAGTCTCTGTTTAAT 3816
1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
3817 AAAGGAATCTGTGTGATGGTGTGCTGCTATCGTTGCACATGTGTGAAGGATTTGTA 3876
1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAla 1320
3877 GGCCTGCATTTGTGAACAGAGACTCAATGAATGCCAGTCAAAACCCATGCTTAAATAATGCA 3936
1321 ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
3937 GTCTGTGAAGACCGGTTGGGGATTTCTTGTGCAAAATGCCACCTGATTTTGGGTACC 3996
1341 ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
3997 CGATGTGGAAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAAATGAGAGCTACCTGT 4056
1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
4057 AAAGCGGTGCCATAGCTTTCAGATGCCCTGTGTGACAGCTGGCTTTCAGAGATCACACTGT 4116

1381 GluLeuAenlleAanGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db GAATTGAACATCAATGAATGTCAGTCTAATCCATGTAGAAATCAGCCACCTGTGTGAT 4176
1401 GluLeuAenSerTyrSerCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db GAATTAATTCATACAGTTGTAATGTTCAGCCAGGATTTTCAGGCAAAAGGTGTGAACA 4236
1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db GAACAGTCTACAGGCTTTAACTCTGATTTTGAAGTTTCTGCGATCTATGATGATGTCATG 4296
1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSer 1460
Db CTAGATGGCATGCTCCCATCTCTCATCTCTCACTCTGATCTCTCTGATGAATTCCTCT 4356
1461 AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db GACGACATGAATATGAACACCAATCTCTATGCACTTGTATGATGATGATGATGATGATGAT 4416
1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db TTGCTCTCTGATGATTAACAGGCTGGCTCTCTTATGTGAATGCGAGGAGGAGGAGGAGGAG 4476
1501 AsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSerAla 1520
Db AACTGTCTCTCGGTGAATGATGGCAGATGGCAGATGATGATGATGATGATGATGATGATGAT 4536
1521 AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyValAlaGlyLeuSer 1540
Db AATGGCATCTGGAAAGTCTATATGATGGGAAATTAATCTGAAGTGGTGGTGGTGGTGGTGGT 4596
1541 ValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLys 1560
Db GTTGGTTTGGCCATACCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4656
1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
Db GGAGGGGATTCACCCAGCTGAGTCTTTTGGGCTCCATGAGCCAGCTCACTCTCTGG 4716
1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db GACTATGTCCTGCTCTCCACAGCAGTGAAGTCACTGGCTACCTCTGCCCCAGAGAACTC 4776
1601 SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db AGTAAAGGAAACGTTAGCATGGCTGATTTCTTGTGAGGAATTTGGGGGAAATGTGAAG 4836
1621 IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db ATCGATCTTAGGACATATTTTGTCTGTATGCTCCACAGCTTAGGAGGTCTAGTGGCTCAT 4896
1641 LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db CTGAGACTGCATCTGAAGATTTAAAGCCAGGTTTCCAAAGTCAATCTGTTCTGTGATCCA 4956
1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
Db GGCTTCCAGCTGGTGGGAAACCCCTGTGCGAGTACTGCTGATCAATCAAGGACAGTGGACACA 5016
1681 ProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPhe 1700
Db CCATCTCTCACTGTGAACGATTTAGCTGTGGGTGGCCACTCTCTTTGGAGAAATGGCTTC 5076
1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db CATTCAGCCGATGACTTCTATGCTGTGGCAGCAGTAACTTACCAGTGCACAAATGGCTAC 5136
1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerThrAsnGlyValSer 1740
Db TATCTATTGGGTGACTCAGGATGTTCTGTACAGATAATGGAGCTGGAAACGGCGTTTCA 5196

1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db CCATCTGCTCTGATGTCGATGAGTGCAGTTCGATGATGATGATGATGATGATGATGATGAT 5256
1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db TGCCTGAACGATAGATGATCTCTACATATGTTTCAATGTCCTCCCGTACACAGGATGGG 5316
1781 LysAsnCysAlaGluProIleLysCysValAlaProGlyAsnProGluAsnGlyHisSer 1800
Db AAAAATCTGTGAGAACCTATATAATGATAGGCTCCAGGAAATCCGAAATGCGCACTCC 5376
1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGlyTyrGln 1820
Db TCAGTGAGATTATACAGTAGTTCGCGAGTCAATTTTCTGTCAGGAGGATACCAAG 5436
1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisLeuIlePro 1840
Db TTGATGGAGTAACCAAAATCAATGTTGGAGTCTGGAGAAATGGAATCATCTAATACCA 5496
1841 TyrCys 1842
Db TATTGT 5502
RESULT 5
US-10-028-248A-7
; Sequence 7, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimketa, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangolli, Esba
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennnda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchervnev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; FILE REFERENCE: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7

; LENGTH: 11152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-028-248A-7

Alignment Scores:

Pred. No.: 0 Length: 11152
Score: 9862.00 Matches: 1781
Percent Similarity: 97.72% Conservative: 16
Best Local Similarity: 96.85% Mismatches: 40
Query Match: 96.51% Indels: 2
DB: 16 Gaps: 1

US-09-977-053-6 (1-1842) x US-10-028-248A-7 (1-11152)

```

Qy 4 ArgLeuAlaPheCysCysThrProGlyLeuAlaLeuValSerGlyTrpAlaThrPheGlnGln 23
Db 83 AGAATTTGGCGGCTTGTGGGCTCGCGCTCGTTTGGGCTGGGCGACCTTTTCAGCAG 142
Qy 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db 143 ATGTCCCGCTCGCGCAATTTTCAGCTTTCGCTCTTCCCGGAGACCGCGCGCGGCGCCCC 202
Qy 44 GlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
Db 203 GGGAGTATCCCGCGCGCGCGCTCTGTGGCGAGCGCGCGGCGGAGCGAGGTGGAGCGG 262
Qy 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
Db 263 CTGGGCCAGGGGTTTC-----CGCGTGGCGCTGTCTCGCGAGCTCAGCGCGCGCTGGAG 316
Qy 84 LeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSerGluLeu 103
Db 317 CTGTGCTTCTCGTGGTGGATGATTCGTCAGCGTGGCGGAGTCAACTTCCGCGAGCGCTC 376
Qy 104 MetPheValArgGlyLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db 377 ATGTTCTGTCGCGAGCTGCTGCTCACTTCCCGTGGTGGCCACGCGCGCGCTGGCC 436
Qy 124 IleValThrPheSerSerIysAsnTyrValValProArgValAspTyrIleSerThrArg 143
Db 437 ATCGTGACCTTCTGCTCAAGAACTACGCTGGTGGTGGCGCGCTCGATTCATCTCCACCG 496
Qy 144 ArgAlaArgGlnHisGlyCysAlaLeuLeuLeuGlnIleProAlaIleSerTyrArg 163
Db 497 CGCGCGCGCGCACAGTGGCGCTGCTCTCCAGAGATCCCTGCCATCTCTCCACCGA 556
Qy 164 GlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
Db 557 GGTGGCGGACCTACACAGGCGCGCTTCCAGCAGCGCGCGCAATTTCTTCTTCATGCT 616
Qy 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAsp 203
Db 617 AGAGAAACTCAACAAAGTTGTATTTCTCATCTCATGGATATTCCAATGGGCGAGAC 676
Qy 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db 677 CCTAGACCAATTTGCAGCGCTACTGCGAGATTGAGAGTGGAGATTTCACTTTTGGCATA 736
Qy 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
Db 737 TGGCAAGGGGACATTCGAGAGCTGATGCATGGCTTCCACCCCAAGGAGGAGCAGCTGT 796
Qy 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeuHisGluAsp 263
Db 797 TACCTGTACACAGTTTGAAGAAATTGAGGCTTTAGGCTTCGCGCGGCGCATTCGATGAAG 856
Qy 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeuCysAspGlu 283
Db 857 CTACCTCTCGGAGGTTTATTCACAGATATATGGTCCACTGCTCATATCTTTGTGATGAG 916
Qy 284 GlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303

```

```

Db 917 GCGAAGGACTGCTGTGACCGGAATGGGAAGCTGCAAAATGTGGGACACACACAGGCCATTTT 976
Qy 304 GluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThrAlaCys 323
Db 977 GAGTGCATCTGTGAAAGGGGTATTACCGGAAGGTCTGCAGTATGAATGCACAGCTTGC 1036
Qy 324 ProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCysIleProCys 343
Db 1037 CCATCGGGGACATACAAACCTGAGCCTCACCAGGAGGATCAGCAGTTCATTCATGTT 1096
Qy 344 ProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysValCysArg 363
Db 1097 CCGGATGAAATACACACTCTCCACTCGGAAGCACATCCCTCGAAGACTGTGTCTGCGA 1156
Qy 364 GluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuValPro 383
Db 1157 GAGGATACAGGGCATCTGGCCAGACCTGTGAACCTTGTCCACTGCGCTGCGCTGAAGCCT 1216
Qy 384 ProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGly 403
Db 1217 CCGAATAATGGTTACTTTATCCAAACACTTGCACACACTTCAATGCAGCCTGTGGG 1276
Qy 404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsn 423
Db 1277 GTCCGATGTCACTCCCTGGATTTGATCTTGTGGGAAGCAGCATCATCTTATGTCTACCA 1336
Qy 424 GlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGln 443
Db 1337 GGTGTTGTGTCGGTTCAGAGAGCTACTGCAGAGTAAGAACATGTCTCATCTCCGCGAG 1396
Qy 444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeu 463
Db 1397 CCGAACAATGGCCACATCAGCTGTTCTACAGGGAATGTATATAGACAAACATGTTTG 1456
Qy 464 ValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGlyAsn 483
Db 1457 GTTGTCTGTGATGAAGGGTACAGACTAGAAAGCAGTGAATGAAGCTTACTTGTCAAGGAA 1516
Qy 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
Db 1517 AGCCAGTGGATGGGCGCAGAACCCCGGTGTGTGGAGCGCCACTGTTCCACCTTTTCAGATG 1576
Qy 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
Db 1577 CCCAAGATGTTCATCATATCCCTCCACCACTGTGGCAAGCAGCAGCCAAATTTGGGACG 1636
Qy 524 IleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
Db 1637 ATCTGCTATGTAAATTTGCGCCAGGGTTTCAATTTATCTGGAGTCAAGAAATGCTGAGA 1696
Qy 544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAspValGlu 563
Db 1697 TGATCACCTTCTGGAAATTTGGAATTCGAGTTCGAGGAGCTGTGTGTAAAGACGTGGAG 1756
Qy 564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnInAspSer 583
Db 1757 GCTCTCAAACTCACTGTCTAAGGACATAGAGGCTTAAGACTCTGGAACAGCAAGATTCT 1816
Qy 584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlyLysValSerVal 603
Db 1817 GCCAATGTTTACCTGGCAGATTCCAAACAGCTAAGCAACCTCTGTGTGAAAGGTGTGATC 1876
Qy 604 HisValHisProAlaPheThrProTyrLeuPheProIleGlyAspValAlaIleVal 623
Db 1877 CCGGTTCATCCAGCTTTTACCCCACTTACCTTTTCCCAATTTGGAAGTGTGTATCGTA 1936
Qy 624 TyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
Db 1937 TACACGCAACTGACCTATCCGCAACAGGCCAGCTGCTGATTTTCCATATCAAGGTTATT 1996
Qy 644 AspAlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGlu 663
Db 1997 GATGCAGAACCACTGCTCATAGACTGGTGCAGATCTCCACCTCCCGTCCAGGCTCGGAG 2056

```


Qy	664	LysValHisAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu	683	Qy	1024	GlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyr	1043
Db	2057	AAAGTATATCCCGAGCTGGAGTACGCTTCTCAGACAACTCAGGGCTGAATG	2116	Db	3137	CAAGATGAAGAGGCAACTTGAGTGCAGAGCTTTGCCCTCTGGATGTACACGGAATAT	3196
Qy	684	ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln	703	Qy	1044	IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr	1063
Db	2117	GTCAATACCAAGAGTCATACACAAAGGAGACCTTTTCCCTCAAGGGAGACTATAGTACAG	2176	Db	3197	ATCCATTCAAGAAACATCTCTGATTGTAAAGCTCAGTGTAAACAGGACCTTACTCATAC	3256
Qy	704	TyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLys	723	Qy	1064	SerGlyLeuGluThrCysGlySerCysProLeuGlyThrTyrGlnProLysPheGlySer	1083
Db	2177	TATACAGCCACTGACCCCTCAGGTAATACAGGATATGTGATATCCATATGTTCATGANA	2236	Db	3257	AGTGGACTTGAGACTTGGAATCGTGCCACTGGCAGCTTATCAGCCAAATTTGGTTCC	3316
Qy	724	GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp	743	Qy	1084	ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle	1103
Db	2237	GGTTCCTCCCTGTGAAATTCATTCACACTGTAAATGGGATTTTATATGCACCTCCAGAT	2296	Db	3317	CGGAGCTGCCTCTCGTGTCCAGAAAAACACCTCAACTGTGAAAAAGAGGACCGGTGAACATT	3376
Qy	744	AsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer	763	Qy	1104	SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys	1123
Db	2297	AATACTGGAGTCAACTGTACATTAACCTTGCTTGGAGGGCTACGATTTCAAGAGGGTCT	2356	Db	3377	TCTGCATGTGGAGTTCCCTTGTCCAGAAAGGAAAAATCTCGCGTTCTGGGTAAATGCCCTGT	3436
Qy	764	ThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrGlu	783	Qy	1124	HisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys	1143
Db	2357	ACTCACAAATATATTCTGTCTTATGAAGATGGCGTCTG3AAACCAACATATACCCTGAA	2416	Db	3437	CACCATGTCTCTGTGACTATTACCAACTAATGCAGGAAGGCTTCTGCTGCCCTGT	3496
Qy	784	TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe	803	Qy	1144	ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPhe	1163
Db	2417	TGCGCAGACTGTGCAAAAACGTTTTGCAAAACCGCGGTTCAAGTCTCTTGAGATGTTC	2476	Db	3497	CCCTTTATGGAACCTACCCCACTCGCTGGTTCAGATCCATCACAGAAATGTTCAAGTTT	3556
Qy	804	TyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu	823	Qy	1164	SerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGlyHisIle	1183
Db	2477	TACAAAGCAGCTCGTGTGTGATGACTCAGATCTGATGAAGAAGTTTTCTGAAGCATTTTGAG	2536	Db	3557	AGTTCAACTTTCTCAGCGGCAGAGAGAAAGTGTGTGTGCCCTCTCTCTTG3ACATATT	3616
Qy	824	ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg	843	Qy	1184	LysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHis	1203
Db	2537	ACGACCTCGGAAAAATGGTCCCATCATTTTTGTAGTGTGACAGAGACATTGACTGCAGA	2596	Db	3617	AAAAGAGGCATGAATCAGCAGTCAGCAAGTCAATGAATGCTCTTTAAACCTTTGCCAC	3676
Qy	844	LeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly	863	Qy	1204	AsnSerGlyThrCysGlnGlnLeuGluArgGlyTyrValCysLeuCysProLeuGlyTyr	1223
Db	2597	CTGAGAGAACCTGACCAAAAATATTTGCTAGANATATATATGACTATGAANAATGGC	2656	Db	3677	AATAGTGGACCTGCCAGCACTTGGGGTGTTATGTTTGTCTCTGTCTCCACTTGGATAT	3736
Qy	864	PheAlaIleGlyProGlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAsp	883	Qy	1224	ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsn	1243
Db	2657	TTTGCAATTGGTCCAGGTGGCTGGGTGACGCTAATAGCTGGATTACTCTTACGATGAC	2716	Db	3737	ACAGGTTTAAAGTGTGAACACAGACATCGATGAGTGCAGCCCACTGCTTGCCTCAACAAT	3796
Qy	884	PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle	903	Qy	1244	GlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGly	1263
Db	2717	TTCTTGACACTGTGCAAGAAACACGACCAAGCATCGCAATCGCAATGCTCTCAGGATT	2776	Db	3797	GGAGTTGTAAAGACCTAGTTTGGGAATTCATTTGTAGTGGCCCATCAGTTTACAGGT	3856
Qy	904	LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer	923	Qy	1264	GlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIle	1283
Db	2777	AAAAGAAGTGCCCATTTATCTGACTATATAAATTAAGTTAATTTTAACTACACAGCTAGT	2836	Db	3857	AAGCACTGTGAATGAACATCAATGAATGTCAGTCTAATCCATGTAGAATCAGGCCACC	3916
Qy	924	ValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeu	943	Qy	1284	CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis	1303
Db	2837	GTGCCATTACCCGATGAAGAAATGATAGCCCTTGAATGGGAAATCAGCAACGACTCCTT	2896	Db	3917	TGTGTGGAATGAATTAATTCATACAGTTGTAAATGTCAGCCAGGATTTTCAGGCAAAAGG	3976
Qy	944	GlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyr	963	Qy	1304	CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu	1323
Db	2897	CAGACATTGGAAATATATACAAATAAATCTGAAAAGGACTCTCTCAACAAAGACCCCATGTAT	2956	Db	3977	TGTGAACAGTATGATATCAACTCAGTGTATTATTAATTAACCTTAATATGTCAGTCTGTGAA	4036
Qy	964	SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLys	983	Qy	1324	AspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGly	1343
Db	2957	TCCTTTTCAGCTTGATCAGAAATATCTATAGCCGACAGCAATTCATTAGAAAACAAAAAG	3016	Db	4037	GACCAGGTTGGGGATCTCTGTGCAAAATGCCACCTGGATTTTGGGTACCCGATGTGGA	4096
Qy	984	AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro	1003	Qy	1344	LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly	1363
Db	3017	GCTTCCCTTCTGCACAGCCAGGCTCAGTGTGAGAGGGCGTATGTGTCAATTTGCCCT	3076	Db	4097	AAGAACGTGATGATGTCTCAGTCAGCATGCAAAAATGGAGCTTACTGTAAAGACGCT	4156
Qy	1004	LeuGlyThrTyrThrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr	1023	Qy	1364	AlaAsnSerPheArgCysLeuAlaAlaGlyPheThrGlySerHisCysGluLeuAsn	1383
Db	3077	TTGGAAACCTTATTAATCTGGAACATTTTCACTGTGAAAGCTGCCGATCGGATCTCTAT	3136	Db	4157	GCCAAATAGCTTCAGGTGCTGTGTGACGTGCTTTCACAGGATCACACTGTAATTAAC	4216
				Qy	1384	IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn	1403

Db 4217 ATCAATGAATGTCAGTCTTAATCCATGTAGAAATCGGCCACTGTGTGGATGAATTAAT 4276
Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluInSer 1423
Db 4277 TCATACAGTGTGTAATGTTCAGCCAGGATTTTCAGSCAAAAGGTGTGAAACAGACAGTCT 4336
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
Db 4337 ACAGGCTTTAACTGGGATTTTGAAGTTTCTGGCATCTATGATATGTCTATGATGGC 4396
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAspMet 1463
Db 4397 ATGCTCCCACTCTCCATGCTCTTAACCTGTAACCTTCTGATGAATCTCTGACGACATG 4456
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
Db 4457 AACTATGACACCACTCTCTATGTCAGTGTATTAACGGCAGCAGCAATACCTTGTCTCTG 4516
Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAlaAsnGlyArgGluLysIleThrAsnCysPro 1503
Db 4517 ACTCAATATTAACGGGTGGGTCTTTATGTGAATGGCAGGAAAGATTAACAACTGTCTCC 4576
Qy 1504 SerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAlaAsnGlyIle 1523
Db 4577 TCGTGATGATGGCAGATGSCATCATATTCATCTGACACAGTACTGTGGAGCC 4636
Qy 1524 TrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu 1543
Db 4637 TGGAGGGTCTATATAAATGGGAAATTAATCTGACGGTGGTACTGGCCTCTCCATTGGCAAA 4696
Qy 1544 ProIleProGlyGlyValAlaLeuValLeuGlyGlnGluGlnAspLysLysGlyGly 1563
Db 4697 GCCATACCTGTGGCGGTGCATTTAGTCTTGGGCAAGAGCAGCAAAAAGAGAGGGG 4756
Qy 1564 PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal 1583
Db 4757 TTCACCGGCTGAGTCTTTTGTGGCTCCATAGCCAGCTCACTCTGGGACTATGTC 4816
Qy 1584 LeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGly 1603
Db 4817 CTGTCTCCACAGCAGGTGAAGTCACTGGCTACTCTCTGCCAGAGGAACTCAGTAAAGGA 4876
Qy 1604 AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLysIleAspSer 1623
Db 4877 AACGTGTAGCATGCTGATTTCTGTGAGAAATGTGGGAAAGTGAAGATCATCTCT 4936
Qy 1624 LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr 1643
Db 4937 AAGAGCATATTTTGTCTGATTTGCCACGCTTGGAGGGTCACTGCTCACTGAGAACT 4996
Qy 1644 AlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln 1663
Db 4997 GCATCTGAAGATTTAAACACAGGTTCCAAAGTCAATCTGTCTGTGAACACCGGCTTCAG 5056
Qy 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuPro 1683
Db 5057 CTGTGGGGAACCTCTGTCAGTACTGTCTGATCAAGACAGTGGACACACCACTCTCC 5116
Qy 1684 HisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAla 1703
Db 5117 CACTGTGAACGCATTCGCTGTGGGGTGCACCTCTTTGGAGAAATGGCTTCAATTCAGCC 5176
Qy 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu 1723
Db 5177 GATGACTTCTATGCTGGGCACACAGTAACTACCACTGCAACCAATGGCTACTATCTATTG 5236
Qy 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys 1743
Db 5237 GGTGACTCAAGGATGTTCTGTACAGTAATGGAGCTGGAACGGGCTTTCACCATCTCTGC 5296
Qy 1744 LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn 1763

Db 5297 TTAGATGTCCGATGAGTGTGCAGTGTGGATCAGATTGTAGTGAGCATGCTCTTTGCCCTGAAC 5356
Qy 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCys 1783
Db 5357 GTAGATGGATCTCTCATATATTTTCATGTGTCCACCGTACACAGGAGATGGGAAAACCTGT 5416
Qy 1784 AlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu 1803
Db 5417 GCAGACCTATAAATGTAAAGCTCCAGGAATCCGGAATCGCAATCTCTCAGGTGAG 5476
Qy 1804 IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly 1823
Db 5477 ATTTATACAGTAGTGTCCGAAAGTCACTTTCTGTCTCAGGAAGGATACCACTTGTGGA 5536
Qy 1824 ValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyrCys 1842
Db 5537 GTAACCAAAATCACAATGTTTGGAGTCTGGAGAAATGGAATCATCTAATACCATTTGT 5593

RESULT 6
US-10-107-782-7
; Sequence 7, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Coleman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zernusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: CuraSeq1 version 0.1
; SEQ ID NO 7
; LENGTH: 11152
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE:
NAME/KEY: CDS
LOCATION: (77) .. (10780)
US-10-107-782-7

Alignment Scores:
Pred. No.: 0 Length: 11152
Score: 9862.00 Matches: 1781
Percent Similarity: 97.72% Conservatives: 16
Best Local Similarity: 96.85% Mismatches: 40
Query Match: 96.51% Indels: 2
DB: 16 Gaps: 1

US-09-977-053-6 (1-1842) x US-10-107-782-7 (1-11152)
QY 4 ArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThrPheGlnGln 23
DB 83 AGAATTGGCGCGCTTGGCTGGGCTCTGGCGCTCTGGCTGGGCTGGGCGACCTTTTCAGCAG 142
QY 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
DB 143 ATGTCCCGCTGGCGCAATTTTCAGCTTCCGCTCTTCCCGGAGCCCGCCGGGCCCCC 202
QY 44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
DB 203 GGGAGTATCCCGCGCGCGCTCTGGCGAGCAAGCGCGGGAGAGTGGAGCGG 262
QY 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuLeuArgGluLeuSerGluArgLeuGlu 83
DB 263 CTGGGCGAGCGGTC-----CGCGTGGCGCTGCTGGCGAGCTCAGCGAGCGCTGGAG 316
QY 84 LeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSerGluLeu 103
DB 317 CTGTGCTTCTGTGGATGATTCGTCCAGCGTGGCGAGTCACTTCCGACGAGCTC 376
QY 104 MetPheValArgGlyLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
DB 377 ATGTTGCTCGCAAGCTGCTGCGAGTCTCCCGTGGTGGCCACCGCGCGGTGGCC 436
QY 124 IleValThrPheSerSerIleValValProArgValAspTrpIleSerThrArg 143
DB 437 ATCGTGACCTTCTCGTCCAGAACTACGTGGTGGTGGCGCGTTCGATTACATCTCCACCCG 496
QY 144 ArgAlaArgGlnHisCysValAlaLeuLeuGlnGlnIleProAlaIleSerTrpArg 163
DB 497 CGCGCGCGCGCAGCACAGTGGCGCTGCTCTCCAGAGATCTCTCCCATCTCTACCGA 556
QY 164 GlyGlyGlyThrThrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
DB 557 GGTGGCGGCACTACACCAAGGCGCTTCCAGCAAGCGCGCAATTTCTTTCATGCT 616
QY 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTrpSerAsnGlyGlyAsp 203
DB 617 AGAGAAAACCTCAACAAAAGTTGTTTCTCATCTGATGATATTTCCAAATGGGCGAGAC 676
QY 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
DB 677 CTAGACCAATTTGACGCGTCACTCGCGAGATTCAGGAGTGGAGATTTCACTTTTGGCATA 736
QY 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
DB 737 TGGCAAGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAGGAGCACTGT 796
QY 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeuHisGluAsp 263
DB 797 TACCTGCTACACAGTTTTCAGAAATTTGAGGCTTTAGCTCGCGGGCAATTCGATGAAGAT 856
QY 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrpLeuCysAspGlu 283
DB 857 CTACCTTCTGGGAGTTTATTCAGATGATATGGTCCACTGCTCATATCTTTGTGATGAG 916
QY 284 GlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303

```

```

DB 917 GGCAGGAGCTGCTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACACAGGCGCATTTT 976
QY 304 GluCysIleCysGluLysGlyTrpTrpGlyGlyLeuGlnTrpGlyCysThrAlaCys 323
DB 977 GAGTGCATCTGTGAAAAGGGGTATTACGGGAAGGCTGTGAGTATGAATGCACAGCTTGC 1036
QY 324 ProSerGlyThrTrpLysProGluGlySerProGlyGlyIleSerSerCysIleProCys 343
DB 1037 CCATCGGGGACATACAACTTGAGGCTCACAGGAGGATTCAGAGTGTGATTCATTCATGT 1096
QY 344 ProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysValCysArg 363
DB 1097 CCGGATGAAAATCACACCTCTCCACCTGGAAGCACATCCCTGGAAGACTGTGTCTGCAGA 1156
QY 364 GluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
DB 1157 GAGGATACAGGGGATCTGGCCAGACTGTGAACATTTGTCCACTGCCCTGCCCTGAGGCT 1216
QY 384 ProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGly 403
DB 1217 CCGGAAAATGTTACTTTATCCAAAACACTTGCACAAACCACTTCANTGCAGCCTGTGGG 1276
QY 404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCysLeuProAsn 423
DB 1277 GTCCGATGTCCACTGGAATTTGATCTTGTGGGAAGCAGCATCATCTTATGTACTCCCAAT 1336
QY 424 GlyLeuTrpSerGlySerGluSerTrpCysArgValArgThrCysProHisLeuArgGln 443
DB 1337 GGTGTTGTGGTCCGGTTTCAGAGACTACTGCAGAGTAGAAGACATGCTCTCATCTCCGCGAG 1396
QY 444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTrpLysThrCysLeu 463
DB 1397 CCGAATACATGCGCCACATCAGCTGTTCTACAGGGAATGTTATATAAGACACATGTTG 1456
QY 464 ValAlaCysAspGluGlyTrpArgLeuGluGlySerAspLysLeuThrCysGlnGlyAsn 483
DB 1457 GTTGCTGTGATCAAGGCTACAGACTAGAGGCGAGTGAAGCTTACTTGTCAAGGAAC 1516
QY 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
DB 1517 AGCAGTGGGATGGGCCAGAACCCCGGTGTGGAGGCCACTGTTCCACTTTCCAGATG 1576
QY 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
DB 1577 OCCAAGATGTATCATATCCCTCCCACTGTGGCAAGCAGCAGCCAAATTTGGGAGC 1636
QY 524 IleCysTrpValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
DB 1637 ATCTGCTATGTAAGTTCGCGCAAGGCTTCATTTTATCTGGAGTCAAAAGAAATGCTGAGA 1696
QY 544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAspValGlu 563
DB 1697 TGTACCACTTCTGGAATAATGGGAATGTGGAGTTTCAGGAGCTGTGTGTAAAGAGCTGGAG 1756
QY 564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnGlnAspSer 583
DB 1757 GCTCTCAATCACTGTCTTAAGGACATAGAGCTTAAGACTCTGGAACAGCAAGATTCT 1816
QY 584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerVal 603
DB 1817 GCCAATGTTACCTGGCAGATTCACAGCTTAAAGACAACTCTGCTGAAAAGGTGTCTAGTC 1876
QY 604 HisValHisProAlaPheThrProTrpTrpLeuPheProIleGlyAspValAlaIleVal 623
DB 1877 CGGTTTCATCCAGCTTTCACCCCACTTACCTTTTCCAAATGGAGATGTGTCTATCGTA 1936
QY 624 TyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
DB 1937 TACACGCACTGACCTATCCGCAACAGGCGCAGCTGCAATTTCCATATCAAGTTATT 1996
QY 644 AspAlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGlu 663
DB 1997 GATGCAGAACCACTGTGTCATAGACTGGTGCAGATCTCCACCTCCGCTCCAGGTCTCGGAG 2056

```

QY 664 LysValHisAlaAlaSerThrAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu 683
DB 2057 AAGGTATCATGCCCAAGCTGGATGAGCTCAGTTCTTCAGACAACTCAGGGGCTGAATTG 2116
QY 684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
DB 2117 GTCAATTACCAAGAGTCATACACAGGAGACCTTTTCCCTCAGGGGAGACTATAGTACAG 2176
QY 704 TyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLys 723
DB 2177 TATACAGCCACTGACCCCTCAGGTAAATAACAGGATATGTATATCCATATTTGTCAAGAA 2236
QY 724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
DB 2237 GGTTCCTCCCTGTGAATTCATTCACACCTGTATAATGGGATTTTATATGACCTCCAGAT 2296
QY 744 AsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
DB 2297 AATCTGGAGTCAACTGTATACATTAACCTTGTGGAGGGCTACGATTTTCAACAGAGGTCT 2356
QY 764 ThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyrThrGlu 783
DB 2357 ACTGACAAGTATTATTGTCTTATGAAGATGGCGTCTCGAAACCAACATATACCCTGAA 2416
QY 784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
DB 2417 TGGCCGAGACTGTGCAAAAAACGTTTTGCAAAACACCGGGTTCAAGTCTCTTGAGATGTC 2476
QY 804 TyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu 823
DB 2477 TACAAAGCACTCTGTGTGATGACTCAGATCTGATGATGAAGAGTTTTCTGAAGCAATTTGAG 2536
QY 824 ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
DB 2537 ACCGACCTGGGAAAAATGGTCCCATCATTTTGTAGTGTATGATGAGAGCAATTTGACGAGA 2596
QY 844 LeuGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly 863
DB 2597 CTGGAGGAGAACCTGACCAAAAAATATTTGCCCTAGAAATATATTAATGACTATGAAAAATGGC 2656
QY 864 PheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAsp 883
DB 2657 TTTTGCAATTTGTTCAGTGTGGTGGGTGCGAGCTAATAGCTGGATTAATCTTACGATGAC 2716
QY 884 PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle 903
DB 2717 TTCCTGCACACTGTGCAAGAAACAGCCACCAAGCATCGCAATGCGCAATGCTCACCGGATT 2776
QY 904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
DB 2777 AAAAGAAAGTGGCCCATTTATCTGACTATAAAATTAAGTTTAATTTTAAACATCACAGCTAGT 2836
QY 924 ValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnArgLeuLeu 943
DB 2837 GTGCCATTACCCGATGAAAGAAATGATACCTTGAATGGGAAATACGCAACGACTCCTT 2896
QY 944 GlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyr 963
DB 2897 CAGACATTTGGAAACTATACAAATATAACTGAAAAAGGACTCTCTCAACAAAGACCCCATGTAT 2956
QY 964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLys 983
DB 2957 TCCTTTTCAGCTTGTGATCAGAAATATCTTATAGCCGACAGCAATTCATTTAGAAACAAANAG 3016
QY 984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003
DB 3017 GCCTTCCCTTCTGCAGACCAAGGCTCAGTGTGTGAGGGCGTATGTGTGTCATTTGCCCT 3076
QY 1004 LeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr 1023
DB 3077 TTGGGAACCTATTATTATCTGGAACATTTACCTGTGAAAGCTGCCGATCGGATCCTAT 3136

QY 1024 GlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyr 1043
DB 3137 CAGATGAGAGAGGGCAACTTTAGTGCAGGCTTTGCCCTCTGGATCTACAGGAATAT 3196
QY 1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
DB 3197 ATCCATTTCAAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACCAAGGACCCCTACTCATAC 3256
QY 1064 SerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySer 1083
DB 3257 AGTGGACTTGAGACTTTGTAATCTGTCCACTTGGGCACTTATCAGCCAAAATTTGGTTCC 3316
QY 1084 ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle 1103
DB 3317 CGAGCTGCTCTCTGTGTCCAGAAAACACCTCAACTGTGTGAAAAGAGGAGCGGTGAACATT 3376
QY 1104 SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys 1123
DB 3377 TCTGCATGTGGAGTTCTCTGTCTCCAGAAAGGAAAATTTCTGCGTTTCTGGGTAAATGCCCTGT 3436
QY 1124 HisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys 1143
DB 3437 CACCCATGCTCTGTGACTATTACCAACCTAATGAGGAAAGGCTTCTGCTGCGCTGT 3496
QY 1144 ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPhe 1163
DB 3497 CCTTTTATGGAACCTACCCCATTCGCTGGTTCCAGATCCATCAGAAATGTTTCAAGTTTT 3556
QY 1164 SerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGlyHisIle 1183
DB 3557 AGTTCAACTTTCTCAGCGGAGAGGAAAGTGTGGTGGCCCTGCTCTCTTGAGCATATT 3616
QY 1184 LysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHis 1203
DB 3617 AAAAAGAGGCATGAATGAATCAGCAGTCAGCAAGTCATGAATGCTTCTTTAACCCCTTGCAC 3676
QY 1204 AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr 1223
DB 3677 AATAGTGGAACTGCCAGCAACTTGGCGTGTATGTTGTCTCTGTCTCCACTTGGATAT 3736
QY 1224 ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsn 1243
DB 3737 ACAGGTTTAAAGTGTGAAACAGACATCGATGAGTGCAGCCCACTGCTTGGCTTCAACAA 3796
QY 1244 GlyValCysLysAspLeuValGlyPheIleCysGluCysProSerGlyTyrThrGly 1263
DB 3797 GGAGTTTGTAAAGACCTTAGTTGGGGAATTCATTTGTGAGTGCCCATCAGTTTACAGGT 3856
QY 1264 GlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIle 1283
DB 3857 AAGCACTGTGAATGAAACATCAATGAATGTCACTTAATCTATCTAGAAATCAGGCCACC 3916
QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
DB 3917 TGTGTGATGAATTAATTAATTCATCAGTTGTAATGTCCAGCCAGGATTTTTCAGCGAAAAGG 3976
QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu 1323
DB 3977 TGTGAACAGTATGATATCACTCAGTGTATTATTAACCTTAATATGACGTCTGTGAA 4036
QY 1324 AspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGly 1343
DB 4037 GACCAGGTTGGGGATTTCTGTGCAAAATGCCACCTGGATTTTGGGTACCCGATGTGGA 4096
QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
DB 4097 AAGAACGTGAGTGTCTCAGTCAGCCATGCAAAAATGGAGCTACCTGTAAAGCGT 4156
QY 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
DB 4157 GCCAATAGCTTCAGGTGCTGTGTGACGTGGCTTCCAGGATCACACTGTGAATTGAAC 4216
QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403

Db 4217 ATCAATGAATGTCAGTCTAATCCATGATAGAAATCAGGCCACCTGTGTGGATGAATTAAT 4276
 Qy 1404 SerTyrSerCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
 Db 4277 TCATACAGTTGTAAATGTTCAGCCAGGATTTTCAGGCCAAAGGTGTGAACAGAACAGTCT 4336
 Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
 Db 4337 ACAGGCTTTAACTCGATTTTGAAGTTTCTGGCATCTATGGATATGTCATGCTAGATGCC 4396
 Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetIysSerSerAspMet 1463
 Db 4397 ATGCTCCCATCTCTCCATGCTCTAACTGCTACCTTCTGGATGAATCTCTGACGACATG 4456
 Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
 Db 4457 AACTATGGAAACACCATCTCTATGCGAGTTGATTAACGGCGGACGACATATCTTCTCTCTG 4516
 Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluIysIleThrAsnCysPro 1503
 Db 4517 ACTGATTATTAACGGTGGTCTTTTATGTCATGCGAGGAGGAAAGATTAACAACTGTCCC 4576
 Qy 1504 SerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAlaAsnGlyIle 1523
 Db 4577 TCGTGAATGATGCGAGATGTCATATATGCAATCACTTGGCAAGTACTGGTGGAGCC 4636
 Qy 1524 TrpIysValTyrIleAspGlyIysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu 1543
 Db 4637 TGGAGGTCTATATAATGGGGAATATATCTGACGGTGGTACTGGCCCTCTCCATTTGGCAAA 4696
 Qy 1544 ProIleProGlyGlyAlaLeuValLeuGlyGlnGluGlnAspIysGlyGlyGly 1563
 Db 4697 GCAATACCTGGTGGGTGCTATGTTCTTGGGCAAGGAGGAGGAGGAGGAGG 4756
 Qy 1564 PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal 1583
 Db 4757 TTCAACCGCGCTGAGTCTTTTGTGGGCTCCATTAAGCCAGCTCAACCTCTGGGACTATGTC 4816
 Qy 1584 LeuSerProGlnGlnValIysSerLeuAlaThrSerCysProGluGluLeuSerIysGly 1603
 Db 4817 CTGTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTGCCAGAGGAACTCAGTAAAGGA 4876
 Qy 1604 AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyIysValIysIleAspSer 1623
 Db 4877 AAGGTGTACATGCGCTGATTTCTGTGAGGAATTTGTGGGAAAGTGAAGATGCTCT 4936
 Qy 1624 LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr 1643
 Db 4937 AAGAGCATATTTTGTCTGATTTGCCACGCTTGGAGGGTCAGTGCCTCATCTGAGAACT 4996
 Qy 1644 AlaSerGluAspLeuIysProGlySerIysValAsnLeuPheCysAspProGlyPheGln 1663
 Db 4997 GCATCTGAAGATTTAAACACCGGTTCCAAAGTCAATCTGTCTGTGAACAGGCTTCCAG 5056
 Qy 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuPro 1683
 Db 5057 CTGGTGGGAACCTGTGCGAGTACTGTCTGAATCAGGACGTGGACACCACTCCCTCCC 5116
 Qy 1684 HisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAla 1703
 Db 5117 CACTGTGAACGCAATCGCTGTGGGGTCCACCTCTTTTGGAGAATGGCTTCCATTCAGCC 5176
 Qy 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu 1723
 Db 5177 GATGACTTCTATCTGCGACACAGTAACCTACAGTGCACCAATGGCTACTATCTATTG 5236
 Qy 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys 1743
 Db 5237 GGTGACTCAGGAGTGTCTGTACAGATTAATGGAGGTGGAAACGGGTTTCCCATCTCTGCG 5296
 Qy 1744 LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn 1763

Db 5297 TTAGATGTCGATGAGTGTGCGAGTTGGATGCGATGATGATGAGCATGCTTCTTGCCTGAAC 5356
 Qy 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyIysAsnCys 1783
 Db 5357 GTAGTGGATCTTACATATGTTTATGTTGTCCTCCACGATACACAGGAGATGGGAAACTGT 5416
 Qy 1784 AlaGluProIleIysCysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu 1803
 Db 5417 GCAGAACCTATAAATGTAAGGCTCCAGGAAATCCGGAATATGCGCACTCTCCTCAGTGGAG 5476
 Qy 1804 IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly 1823
 Db 5477 ATTATACAGTAGGTGCGGAGTCAATTTTCGTGTCAGGAAGGATACAGTTGATGGGA 5536
 Qy 1824 ValThrIysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyrCys 1842
 Db 5537 GTAACCAAAATCAGATTTTGGAGTCTGGAGATGGATCATCTAATACCATATTTGT 5593

RESULT 7

US-10-028-248A-5
 ; Sequence 5, Application US/10028248A
 ; Publication No. US20030235882A1

GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Malyankar, Uriel
 ; APPLICANT: Shenoy, Suresh
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Gangolli, Bha
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Li, Li
 ; APPLICANT: Taupier Jr, Raymond J
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Zerhusen, Bryan
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Colman, Steven
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Si, Jingsheng
 ; APPLICANT: Edinger, Shlomit
 ; APPLICANT: Stone, David
 ; APPLICANT: Sciore, Paul
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Rothenberg, Mark

; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods (

FILE REFERENCES: 21402-222

CURRENT APPLICATION NUMBER: US/10/028,248A

PRIOR FILING DATE: 2001-12-19

PRIOR APPLICATION NUMBER: 60/256619

PRIOR FILING DATE: 2000-12-19

PRIOR APPLICATION NUMBER: 60/262959

PRIOR FILING DATE: 2001-01-19

PRIOR APPLICATION NUMBER: 60/272408

PRIOR FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: 60/285189

PRIOR FILING DATE: 2001-04-20

PRIOR APPLICATION NUMBER: 60/308039

PRIOR FILING DATE: 2001-07-26

PRIOR APPLICATION NUMBER: 60/311266

PRIOR FILING DATE: 2001-08-09

NUMBER OF SEQ ID NOS: 211

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 11158

TYPE: DNA

ORGANISM: Homo sapiens

US-10-028-248A-5

Alignment Scores:

Db 2117 GAATTTGGTCATTACAGAGTATACACAGAGAGACCTTTTCCCTCAAGGGGAGACTATA 2176
Qy ValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
Db 2177 GTACAGTATACAGGCACCTGACCCCTCAGGCAATAACAGGACATCTGTATATCCATATTTGTC 2236
Qy IleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
Db 2237 ATAAAGGTTCTCCCTGTGAAATTTCCATTTCAACCTGTAAATGGGGATTTTATATGCACT 2296
Qy ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
Db 2297 CCAGTATATCTGAGTCACTGATCAATTAATCTTCTGGAGGGCTATGATTTACAGAA 2356
Qy GlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThr 781
Db 2357 GGGTCTACTGACAGTATATTTGTGCTTATGAGATGGCTTTTGCAAACCCAGGGTTCAAGTCTTTGAG 2416
Qy ThrGluTrpProAspCysAlaLysIysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
Db 2417 ACTGAATGGCCAGACTGTGCCAGTAAGCGTTTTCGAAACCCAGCGGTTCAGTCTTTGAG 2476
Qy MetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAla 821
Db 2477 ATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAGAGAGTTTCTGAGCA 2536
Qy PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
Db 2537 TTTGAGACGACCTCGGGAATAATGGTCCCATCAITTTGTAGTGATGACAGAGCAATTGAC 2596
Qy CysArgLeuGluGluAsnLeuThrLysIysTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
Db 2597 TGCAGACTGGAGGAGACCTGACCAAAAAATATTTGCCGTAGAAATATAATATGACTATGAA 2656
Qy AsnGlyPheAlaIleGlyProGlyTyrGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyr 881
Db 2657 AATGGCTTTGCATTTGGTCCAGTGGCTGGGTGACGCTAATAGCTTACTCTTAC 2716
Qy AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
Db 2717 GATGACTTCTCGACACATGTGCAAGAAACAGCCACAGCATCGGCAATGCCAAGTCCTCA 2776
Qy ArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThr 921
Db 2777 CGGATTTAAAGAGAGTGGCCCATTAATCTGACTATATAAATTAAGTTAAATTTTAAACATACA 2836
Qy AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnArg 941
Db 2837 GCTAGTGTGCCATTTACCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAACGA 2896
Qy LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961
Db 2897 CTCCTTCAGACATTTGGAACATATCAAAATAAATCTGAAAGGAGCTCTCAACAAAGACCC 2956
Qy MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
Db 2957 ATGTATTTCTTTTACGTTCATCAGTAATAATCTTATAGCCGACAGCAATTCATTAGAAACA 3016
Qy LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
Db 3017 AAAAAGGCTTCCCTTCTGACAGCAGGCTCAGTGTGAGAGGGCGTATGTGTGTCAAT 3076
Qy CysProLeuGlyThrTyrThrAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
Db 3077 TGCCCTTTGGCAACCTATTAATCTGCAACATTTTCACTTCACTGTAAGAGCTGCCGATCGGA 3136
Qy SerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyrThr 1041
Db 3137 TCCTATCAAGATGAAGAGGCAACTTGAAGTGAAGCTTTTGCCCTCTGGGATGTACAG 3196
Qy GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061

Db 3197 GAATATATCCATTCAAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAGGCACCTAC 3256
Qy SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPhe 1081
Db 3257 TCATACAGTGTGACTTGAAGATCTGTGTCACCTGGGCACCTTATCAGGCCAAATTT 3316
Qy GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaVal 1101
Db 3317 GGTTCGGGAGCTGCTCTCTGTCAGAAACACCTCAACTGTGAAAAGAGAGCGGTG 3376
Qy AsnIleSerAlaCysGlyValProCysProGluLysPheSerArgSerGlyLeuMet 1121
Db 3377 AACATTTCTGCAATGTGAGTTCCTTGTCCAGAAAGGAAATTTCTCGCGTTCTCGGTTAATG 3436
Qy ProCysHisProCysProAspAspTyrTyrGlnProAsnAlaGlyValAlaPheCysLeu 1141
Db 3437 CCCTGTCAACCATGTCTCTGTCGATTAATCAACCTAATGACAGGAGGCCCTTTCGCTG 3496
Qy AlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
Db 3497 GCCTGTCCCTTTTATGGAACCTACCCCATCTGCTGGTCCAGATCCATCAGAAATGTTC 3556
Qy SerPheSerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGly 1181
Db 3557 AGTTTATAGTTCAACTTTCTCAGCGCAGAGGAAAGTGTGGTGGCCCTGCTCTCTGGA 3616
Qy HisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro 1201
Db 3617 CATATTTAAAGAGGCAATGAATCAGCAGTCAAGCAAGTCATGAATGCTTTTAAACCT 3676
Qy CysHisAsnSerGlyThrCysGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
Db 3677 TGCCACATAGTGAACCTGCGACACTTGGCGCTGGTTATGTTTGTCTCTGTCCACTT 3736
Qy GlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
Db 3737 GGATATACAGGTTTAAAGTGTGAACAGACATCGATGAGTGCAGGCCACTGCTTGGCTC 3796
Qy AsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
Db 3797 AACATGGAGTTTGTAAAGACCTAGTTGGGAAATTCATTTGTGAGTGCCCATCAGGTAC 3856
Qy ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerSerProCysLeuAsnLys 1281
Db 3857 ACAGTAAAGCACTGTGAATTCGAACATCAATGAATGTCAGTCTAATCCATGTAGAATCAG 3916
Qy GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301
Db 3917 GCCACCTGTGTGGATGAATTAATTCATACAGTTGTAATTCAGCCAGGATTTTCAGCG 3976
Qy LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
Db 3977 AAAAGGTTGAAACAGGATGATGATCAACTCAGTGTATTATAATTAACCTTAATTAATGAGTC 4036
Qy CysGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArg 1341
Db 4037 TGTGAGACAGGTTGGGGGATTTCTTGTGCAATGCCCACTGGATTTTGGGTACCGA 4096
Qy CysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLys 1361
Db 4097 TGTGGAAGAACGTCGATGAGTCTCAGTCAGCCATGCAAAAATGAGAGTACCTGTGTAA 4156
Qy AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
Db 4157 GACGCTGCCAATAGCTTCAGGTGCTGTGTGCGCTGGCTTCACAGGATCAGCTGTGAA 4216
Qy LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
Db 4217 TTGAACATCAATGAATGTCACTAAATCCATGATGAATCAGGCCACCTGTGTGATGAA 4276
Qy LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
Db 4277 TTAATTTTATACAGTGTGTAATGTTCAGCAGGATTTTCAGGCAAAAGGTGTGAAACAGAA 4336

QY 1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrValMetLeu 1441
 DB 4337 CAGTCTACAGGCTTTAACTGATTTGAAGTTCTGGCATCTATGGATATGTCATGCTA 4396
 QY 1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAsp 1461
 DB 4397 GATGGCATGCTCCCATCTCTCATGCTCTCACTGCTCACTCTCTGGATGNAATCTCTGAC 4456
 QY 1462 AspMetAsnTyrGlyThrProLysSerTyrAlaValAspAsnGlySerAspAsnThrLeu 1481
 DB 4457 GACATGAATATGGAACACCAATCTCTATGTCAGTTGATACGGCAGCAGCAATACCTTG 4516
 QY 1482 LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyValArgGluLysIleThrAsn 1501
 DB 4517 CTCCTGACTGATATAAAGGGTGGGTCTTTATGTAATGGCAGGAGAAAGATAACAAC 4576
 QY 1502 CysProSerValAsnAspGlyArgTyrHisHisIleAlaIleThrTyrThrSerAlaAsn 1521
 DB 4577 TGTCTCTCGGTGAATGATGGCAGATGGCATCATATGTCATCTTGGCAGAGTACTGCT 4636
 QY 1522 GlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerVal 1541
 DB 4637 GGAGCTGGAGGCTCTATATAAATGGGAATTTATCTGACGGTGGTACTGGCTCTCCATT 4696
 QY 1542 GlyLeuProLysProGlyGlyGlyAlaLeuValLeuGlyGlnGluAspLysGly 1561
 DB 4697 GGCMAAGCCATACCTGGTGGCGGTGATTTGTTGGGCAAGAGCAAGCAAAAAGGA 4756
 QY 1562 GluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAsp 1581
 DB 4757 GAGGGGTTCAACCCCGCTGAGTCTTTGTGGCTCCATTAAGCCAGCTCAACTCTGGGAC 4816
 QY 1582 TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer 1601
 DB 4817 TATGTCTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTGCTGCCAGGAGAACTCACT 4876
 QY 1602 LysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyValLysIle 1621
 DB 4877 AAAGGAACCGTGTATGACATGGCTGATTTCTGTTCAGGAATTTGTGGGAAAGTGAAGATC 4936
 QY 1622 AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeu 1641
 DB 4937 GATTTCAAGACATATTTGTCTGATGCCCAAGCTTGGAGGGTCACTGGCTCATCTG 4996
 QY 1642 ArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly 1661
 DB 4997 AGAAGCTGCATCTGAAGATTTAAACCCAGGTTCAGAGTCAATCTGTTCTGTAACCAAGC 5056
 QY 1662 PheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnPro 1681
 DB 5057 TTCCAGCTGGTGGGAACCTGTGTCAGTACTGTCTGAATCAAGGACAGTGGACACAACA 5116
 QY 1682 LeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHis 1701
 DB 5117 CTCCTCCACTGTGACGCAATTCCTGTGGGGTGGCCACTCTCTTGGAGATGGCTTCAT 5176
 QY 1702 SerAlaAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyr 1721
 DB 5177 TCAGCGATGATCTTCTATGCTGGCAGCAGATTAACCTACAGTGAACAATGGCTACTAT 5236
 QY 1722 LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerPro 1741
 DB 5237 CTATTGGGTGATCTCAAGGATTTCTGTACAGATAATGGAGCTGGAAACGGCGTTTCACCA 5296
 QY 1742 SerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCys 1761
 DB 5297 TCTGCTTGATAGTGTCAAGTGTGTCAGTTGGATTCAGATTTGATGAGCATGCTCTCTGC 5356
 QY 1762 LeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLys 1781
 DB 5357 CTGAACGTAGATGATCTCTACATATGTTCTATGTGTCCCAAGTACACAGGAGATGGGAAA 5416

QY 1782 AsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSer 1801
 DB 5417 AACTGTGCAGAACCTATTAATGTAAGGCTCCAGGAATCCGGAATATGCCACTCTCTCA 5476
 QY 1802 GlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeu 1821
 DB 5477 GGTGAGATTTATACAGTAGTGGTGGCCGAAGTCACATTTTCGTGTGAGGAAGATACCAATTG 5536
 QY 1822 MetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisAlaLeuIleProTyr 1841
 DB 5537 ATGGGAGTAACCAAAATCACATGTTTGGAGTCTGGGAATGAATCATCTAATACCATAT 5596
 QY 1842 Cys 1842
 DB 5597 TGT 5599

RESULT 8
 US-10-107-782-5
 ; Sequence 5, Application US/10107782
 ; Publication No. US20040018970A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Boldog, Perenc,
 ; APPLICANT: Casman, Stacie
 ; APPLICANT: Colman, Steve,
 ; APPLICANT: Edinger, Shlomit,
 ; APPLICANT: Gangolli, Esha,
 ; APPLICANT: Kekuda, Ramesh,
 ; APPLICANT: Li, Li,
 ; APPLICANT: Liu, Xiaohong,
 ; APPLICANT: Malyankar, Uriel,
 ; APPLICANT: Miller, Charles,
 ; APPLICANT: Millet, Isabelle,
 ; APPLICANT: Patturajan, Meera,
 ; APPLICANT: Rothenberg, Mark,
 ; APPLICANT: Sciore, Paul,
 ; APPLICANT: Shency, Suresh,
 ; APPLICANT: Shinkets, Richard,
 ; APPLICANT: Si, Jingsheng,
 ; APPLICANT: Smithson, Glenda,
 ; APPLICANT: Spytek, Kimberly,
 ; APPLICANT: Stone, David,
 ; APPLICANT: Taupier, Raymond, jr.,
 ; APPLICANT: Tchernev, Velizar,
 ; APPLICANT: Vernet, Corine,
 ; APPLICANT: Zerhusen, Brian
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
 ; FILE REFERENCE: 21402-222CIP
 ; CURRENT APPLICATION NUMBER: US/10/107,782
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: 10/028,248
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,619
 ; PRIOR FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 60/262,959
 ; PRIOR FILING DATE: 2001-01-19
 ; PRIOR APPLICATION NUMBER: 60/272,408
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/285,189
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: 60/308,039
 ; PRIOR FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: 60/311,266
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/279,344
 ; PRIOR FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 215
 ; SOFTWARE: CuraseqList version 0.1
 ; SEQ ID NO 5
 ; LENGTH: 11158
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS

Db 2057 TCGGAGAGGTACATGCCGCAAGCTGGGATGAGCTCTCTCAGACAATCTCAGGTGCT 2116
Qy 682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIle 701
Db 2117 GAATGGTTCATTACAGAGGTATACACAGAGAGACCTTTCCCTCAAGSGGAGACTATA 2176
Qy 702 ValGlnIleThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
Db 2177 GTACAGTATACAGGCACCTGACCCCTCAGGCAATACAGAGACATGTGATATCCATATTGTC 2236
Qy 722 IleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
Db 2237 ATAAAAAGGTCTCCCTGTGAATTCATTCACACCTGTAAATGGGGATTTTATATGCACT 2296
Qy 742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
Db 2297 CCAGATATACCTGGAGTCACTGATACATTAACTTCTGGAGGCTATGATTTTCACAGAA 2356
Qy 762 GlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrProThrTyrThr 781
Db 2357 GGGTCTACTGCAAGTATATTGTGCTTATGAAGATGGCGTCTGGAAACCAACATATACC 2416
Qy 782 ThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
Db 2417 ACTGAATGGCCAGACTGTGCCAGTAGCGGTATTGCAAAACCACGGGTTCAGTCCCTTTGAG 2476
Qy 802 MetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAla 821
Db 2477 ATGTTCTCAAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAGTTTCTGAAACA 2536
Qy 822 PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
Db 2537 TTTGAGACGACCTCGGAAAAATGGTCCCATCAATTTGTAGTGAATGCGAGGACATTGAC 2596
Qy 842 CysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
Db 2597 TGCAGACTGGAGGAGAACCTGACCAAAATAATTCCTAGATAATAATTAATGACTATGAA 2656
Qy 862 AsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyr 881
Db 2657 AATGGCTTTGCAATTTGTCTCAGTGGCTGGGTGCGAGCTAATPAGCTGGATCTTACT 2716
Qy 882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
Db 2717 GATGACTTCTCTGGACACTGTGCAAGAAACAGCCACAGCATCGGCAATGCCAAGTCTCA 2776
Qy 902 ArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThr 921
Db 2777 CGATTAAAGAGTGGCCCATTAICTGACTATAAATTAAGTAAATTTTAAACATCACA 2836
Qy 922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrProGluAsnGlnArg 941
Db 2837 GCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAAAATCAGCAACGA 2896
Qy 942 LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961
Db 2897 CTCCTTCAGACATGGAACTATATCAATAAATAACATGAAGAGGACTCTCAACAAAGCCCC 2956
Qy 962 MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
Db 2957 ATGTATTCCTTCAGCTTCATCAGAAATATTATAGCCGACAGCAATTCATTAGAAACA 3016
Qy 982 LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
Db 3017 AAAAAGGCTTCCCTTCTGCAGACAGGCTCAGTGTCTGAGAGGGCGTATGTGTCAAT 3076
Qy 1002 CysProLeuGlyThrTyrTrpAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
Db 3077 TGCCCTTTGGGAACCTATTATTAATCTGGAAACATTTCACTGTGGAAGCTGCCGGATCGGA 3136
Qy 1022 SerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyrThr 1041

Db 3137 TCCTATCAAGATGAAGAGGGCAACTCTGAGTGCAAGCTTTTGCCCTCTGGGATGTACACG 3196
Qy 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
Db 3197 GAATATATCCATTCAGAAACATCTCTGATTTGAAAGCTCAGTGTAAACAGAGCACCTAC 3256
Qy 1062 SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPhe 1081
Db 3257 TCATACAGTGGACTTGAGACTTGATGCTGCTCAGCTGGGACCTTTATCAGGCNAATTT 3316
Qy 1082 GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaVal 1101
Db 3317 GGTTCGCGAGCTGCTCTCGTGTCCAGAAAAACCTCAACTGTGAAAAAGAGAGCGGTG 3376
Qy 1102 AsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMet 1121
Db 3377 AACATTTCTGCATGGAGTTCTCTGTCCAGAAAGAAAAATTCGCGTCTCGGTAAAG 3436
Qy 1122 ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeu 1141
Db 3437 CCTGTCAACCATGCTCTCGTGAATTTACCACTTAATGAGGAGAAAGCCCTTCTGCTG 3496
Qy 1142 AlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
Db 3497 GCCTGTCTCTTTATGGAACCTACCCCAATTCGCTGTTCCAGATCCATCACAGATGTTCA 3556
Qy 1162 SerPheSerSerThrPheSerAlaAlaGluGluSerValProProIleSerLeuGly 1181
Db 3557 AGTTTTAGTTCAACTTTCTCAGCGCAGAGGAAGTGTGTGCTGCCCTGCTCTCTTGA 3616
Qy 1182 HisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro 1201
Db 3617 CATATTAAGAGAGCATGTAATTCAGCAGTCAAGCAAGTCATGAATGCTCTTTAACCT 3676
Qy 1202 CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
Db 3677 TGCCACAATAGTGAACCTGCGCAGCAACTTGGCGCTGGTATGTTGTCTCTGTCCACTT 3736
Qy 1222 GlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProleuProCysLeu 1241
Db 3737 GGATATACAGGTTTAAAGTGTGAAACAGACATCGATGAGTGCAGGCCACTGCTTCCCTC 3796
Qy 1242 AsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
Db 3797 AACATGAGTTTGTAAAGACCTAGTTGGGAAATTCATTTGTGAGTGGCCATCAGGTTAC 3856
Qy 1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLys 1281
Db 3857 ACAGGTAAGCACTGTGAATTTGAACATCAATGAATGTCACTCTAATCCATGATAGAAATCAG 3916
Qy 1282 GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301
Db 3917 GCCACTGTGTGGATGAATTAATTCATACAGTTGTAAATGTGAGCCAGCAGGATTTTCAGGC 3976
Qy 1302 LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
Db 3977 AAAAGGTGTGAACAGATATGATCACTCAGTGTATTATTAATCACTTAATATGCAATC 4036
Qy 1322 CysGluAspGlnValGlyGlyPheLysCysLysCysProProGlyPheLysGlyThrArg 1341
Db 4037 TGTGAAGACCAAGTTGGGGGATTTCTGTGCAAAATGCCACCTGGATTTTGGGTACCCGA 4096
Qy 1342 CysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLys 1361
Db 4097 TGTGAAGAGAGCTCGATGAGTGTCTCAGTCAGCCATGCAAAATGAGACTACTCTGAA 4156
Qy 1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
Db 4157 GAAGTGGCAATAGCTTTCAGGTGCTGTGTGAGTGGCTTTCAGAGGATCACACTGTGAA 4216
Qy 1382 LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
Db 4217 TTGAACATCAATGAATGTCACTTAATCCATGTAGAATCAGGCCACCTGTGTGGATGAA 4276

476 CTCACGAGCTCAAGTTCGTGGCGAAGCTGCTGTCGAGCTTCCCGTGGTGTCCACGGCC 535
120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
536 AGCGTGTGGCATCGTCACTTCTCATCCAAAGAACACGCTGTGGTGGCGGTGATAC 595
140 IleSerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluLeuProAla 159
596 ATCTCCACGACCGCGCGCACCAACACAGTGTGGCTGCTCAGCGCGGAGATCCCGGCC 655
160 IleSerTyrArgGlyGlyGlyThrThrLysGlyAlaPheGlnGlnAlaGlnIle 179
656 ATCACTACCGCGGTGGTGGACCTATACCAAGGGCGCTTCACAGACCGCGCAATC 715
180 LeuLeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSer 199
716 CTTCGTCACTCTAGAGAAACTCCACCAAGTCATATTTCTCATCCGCGGCTATTCC 775
200 AsnGlyGlyAspProArgProIleAlaLeuSerLeuArgAspSerGlyValGluIlePhe 219
776 AATGGCGGAGACCCAGACCTATTTCGAGCATCGCTTCGGGATTTCCGAGTGGAGATCTTC 835
220 ThrPheGlyIleTrrGlnGlnAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
836 AGTTTCGGGATTTGGCAGGGGATATCCGGGAATTCGAGATGATGATGCTTCCACCGGAG 895
240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
896 GAAGAACATTGTACTGCTCCACAGTTTGAAGAAATTTAGGCTTTAGCTCCGAGGCG 955
260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
956 TTGCATGAAGATCTACTTCTCGGAGTTTATCCAGAGGATATGCGCCACTGCTCTTAT 1015
280 LeuCysAspGluGlyLysAspCysCysAspAlaGlnMetGlySerCysLysCysGlyThrHis 299
1016 CTCGTGAGGCTGGGAAGACTGCTGTGACAGATGGCCAGCTGCCTCAATGTGGACACAC 1075
300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlySerGlyLeuGlnTyrGlu 319
1076 ACGGTCAATTTGAATGCATCTGTGAGAGGGCTATTACGGGAAGGTCTGCAGCATGAG 1135
320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
1136 TGCACAGCTTGGCCATCAGGACATATAAGCCGAGCTTCTCCAGGAGGAATCAGCAC 1195
340 CysIleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAsp 359
1196 TGCATCCCATGCTCGACGTAGCCACACTCCCACTGGAGACACTTCCCTGTGAGAC 1255
360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
1256 TGGGTGTGGAGGGGATACAGAGATCTGGCCAGACCTGTGAGGTGTGCTGCT 1315
380 AlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
1316 GGCCTGAAGCCTCCTGAAATGTTTTTTTATACAAACACTTGCAGAAACTACTTCAAT 1375
400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu 419
1376 GCGGCTGTGGGTTCGATGTGCGCGGCTTGACCTTGTGGAGAGCAGCATCATTTG 1435
420 CysLeuProAsnGlyLeuTrrPrrSerGlySerGluSerTyrCysArgValArgThrCysPro 439
1436 TGTCAACCAATGTTGTGTTCTGGGACAGAAAGCTTTCGACAGATGAGAACGTCGCC 1495
440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys 459
1496 CACCTCCGACAGCCCAACAGGCCACATCAGCTGCTTCACTGCGGAAATGCTCTCAAC 1555
460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr 479
1556 ACCCTGTGTTGGTTACCTTCGATGAGGATACAGATTAGAGGACAGCTAGGCTTACC 1615

480 CysGlnGlyAsnSerGlnTrrPrrAspGlyProGluProArgCysValGluArgHisCysSer 499
1616 TGTCAAGGAATGCCCAGTGGGATGGCCAGAGCCCGGTGTGTAGAACGCCATTTGCC 1675
500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
1676 ACCTTCCAGAGCCCAAGAGGCTCATCTTCTCCACCCGCTGGCGAAGCAGCCCGCC 1735
520 LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
1736 AGGCTGGGATGACCTGTACGCTAAGCTCCCGCAGGATACATTTTATCCGCGGTGAGA 1795
540 GluMetLeuArgCysThrThrSerGlyLysTrrPrrAsnValGlyValGlnAlaValCys 559
1796 GAA---GTGAGATGTGCCACATCTGGAGTGGAGTGCCAAAGTTTCAGACAGCTGTGTGC 1852
560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu 579
1853 AAGATGTGGAGGCTCCCAAAATCAGCTGTCCAATGACATTTGAGGCAAGACTCGGGAG 1912
580 GlnGlnAspSerAlaAsnValThrTrrPrrGlnIleProThrAlaLysAspAsnSerGlyGlu 599
1913 CAGCAGGACTCTGCTAATGCTGGCAAGTCCCAACAGCTAAAGACAACTCTCGTGAA 1972
600 LysValSerValHisValHisProAlaPheThrProTrrPrrLeuPheProIleGlyAsp 619
1973 AAGGTGTAGTCCAGCTCCACCCAGCTTTACCCCACTTACCTTTCCTCAATTTGAGAC 2032
620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
2033 GTGGCATCACTACACGCAACCACTCATCGGTAAACAGCAGCTGCACTTTCTAC 2092
640 IleLysValIleAspAlaGluProProValIleAspTrrPrrCysArgSerProProVal 659
2093 ATTAAGTTCATTGATGTGGAA CGGCTGTATAGATTGGTGGCGATCTCCACCTCCAATC 2152
660 GlnValSerGluLysValHisAlaAlaSerTrrPrrAspGluProGlnPheSerAspAsnSer 679
2153 CAGGTGTAGAGAGGAGCAGCCCTGCAGCTGGGATGGGCTCAGTTCTCAGACAACTCC 2212
680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
2213 GGGCTGAATTTGTCATTTACAGCAGCTCACACAGGGGACATGTTTCTCATGGGAA 2272
700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
2273 ACGGTGTGTGTACACAGCCTGACCCCTCAGGCAACAAACAGGACCTGTGACATCCAC 2332
720 IleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739
2333 ATTTGTCATAAAGTTTCTCCTGTGAGGTCCCTTCACCCTGTAAACGGGAGCTTTATC 2392
740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
2393 TGTGCCAGGATAGTGTGAGTTACTGTAGCTGAGCTGCAAGGAGGCTATGATTTC 2452
760 ThrGluGlySerThrAspLysTrrPrrCysAlaTyrGluAspGlyValTrrPrrProThr 779
2453 ACAGAGGCTCACCTGAGAGTACTACTGTGCTTTTGAAGATGATGATCTCGGAGACCA 2512
780 ThrThrThrGluTrrPrrAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSer 799
2513 TACTCTACAGATGGCCAGACTGTGCTATAAACGTTTTTGAACCACTGTTTCAAGTTC 2572
800 PheGluMetPheTrrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
2573 TTTGAATGCTATACAAACCACTGCTGTGTGATGACATGATCTGTTTAAAGAGTTTCT 2632
820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
2633 GCAGCATTTGAGACTACCTCGGGAAACATGGTCCCGTCTTTTGTATACGATGCTGATGAC 2692

840 IleAspCysArgLeuGluGluAanLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
Db
2693 ATTGACTGCAAGACTGGAGGAC---CTGACCAAAATTAATCTGCATCGATGATATATTACAC 2749
Qy
860 TyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrPheGlyAlaAlaAanArgLeuAspTyr 879
Db
2750 TATGAAATGGCTTTGCAATTTGGACCAAGGAGGCTGGGGTGCGAGCAACAGGCTGGATTAT 2809
Qy
880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAanAlaLys 899
Db
2810 TCTACGATCACTTCTCGATGTTGTACAGGAACACCCACCGATGTGGCAAGGCCAGA 2869
Qy
900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn 919
Db
2870 TCTCAGGATTAAGAAGACTGCTCCCATTTCTGACCCCAAAATTCAGCTAAATTTTAC 2929
Qy
920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrPheGluAsnGln 939
Db
2930 ATCACAGTAGCTGCTCCACTCCAGAGGAAGAAACGATACCTTGAATTTGGAGAAATCAG 2989
Qy
940 GlnArgLeuGluGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959
Db
2990 CAGCGACTCATTAAGACTTTGGAACCAATCACCAATGCGCTGAAAAGACACTTTGAATAAA 3049
Qy
960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
Db
3050 GAGCCCATGATTAATTTTCAGCTCGCTCGGAACAGTGTGTGCTGACGAAATTCCTC 3109
Qy
980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
Db
3110 GAAACAGAAAGGCTTTCTCTCTTCAGACACAGGCTCTGTGTGAGGGGGCGCATGTGT 3169
Qy
1000 ValAsnCysProLeuGlyThrTyrThrAsnLeuGluHisPheThrCysGluSerCysArg 1019
Db
3170 GTCAACTGCCCTCGGGAACCTCTTACTCTCTGGAGCATTCACACTGTGGAAGCTGCTC 3229
Qy
1020 IleGlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
Db
3230 ATGGGATCTTACCAAGATGAAGAGGGCAGCTGGATGCAAGCTCTGTCTCCCAAGGACT 3289
Qy
1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
Db
3290 CACCGGGAATACCTCCATTCAGAAGCGTCTCTGAATGCAAGCTCAGTGAAGCAAGGC 3349
Qy
1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
Db
3350 ACCTACTCTTCCAGTGGGCTGGAGACTCGGAATGGTGTCCGCTGGGATCTTATCAACCG 3409
Qy
1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
Db
3410 GAATTTGGATCCCGAGCTGCTCTCTATGCCCAGAAACCAACACGCGGTGAAAGAGGA 3469
Qy
1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
Db
3470 GCGGTGCACATCTCTGCTGTGGAGTGCCCTGCCAGTAGAGAAATCTCCCGTTCTGGG 3529
Qy
1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
Db
3530 CTAAACCCCTGCTACCTTTGCCCTCGAGACTATTACCAACCAATGAGGGAAGTCTTC 3589
Qy
1140 CysLeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGlu 1159
Db
3590 TGGCTGCTTGTCTCTTTATGGAATCTACCAACCATCACTGGCGCCAGCTCCATCAGAC 3649
Qy
1160 CysSerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSer 1179
Db
3650 TGCTCAAGTTTATAGCTCTACTTTCTCAGCAGCAGAAAGACATAGTGCCTCGTGCC 3709
Qy
1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe 1199
Db
3710 CCTGGACATCCCAAGACATGATGAGTACGAGTCAAGTCTTTTCAAGAAATCTCTT 3769
Qy
1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCys 1219

3770 AACCCCTGGCCACCAAGTGGAACTCTGCCAAGCTGGGGCGTGGTATGTCTGTCTCTGC 3829
Qy
1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
Db
3830 CCACCTGGATACACAGCTTAAAGTGTGAACACAGATATTGAATGAATGACGCTCTCTGCCT 3889
Qy
1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
Db
3890 TGCTCATGGTGGATTTGTAGAGACCAAGTTGGGGGATTCAGTGGGAATGTTTCATTG 3949
Qy
1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu 1279
Db
3950 GGCTATTTCAGTCAAAATATGTGAAGAAATATAAATGAGTGTATCTCCAGCCCTTGTCTTA 4009
Qy
1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
Db
4010 AATAAAGGAACCTGCACTGACCGCTTGGCAAGCTACCGCTGTACTGTGTGAAGGATAC 4069
Qy
1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
Db
4070 ATGGGTGTGACCTGTGAAACAGAGCTCAATGAATGCCAGTCAAGCCCTGCTTAAACAC 4129
Qy
1320 AlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGly 1339
Db
4130 GCAGTTTGTAAAGACCAAGTTGGGGGTCTCTCGTGCAATGCCCAACCCGGATTTTGGGT 4189
Qy
1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr 1359
Db
4190 ACTCGGTGTGAAAAAATGTGGATGAGTGTCTCAGTCAGCCATGCCAAAAATGGAGCCACT 4249
Qy
1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
Db
4250 TGTAAAGTGTGCCAACACAGCTTCAGGTGTCAATGTCCAGCAGGCTTCCAGAGGACACAC 4309
Qy
1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
Db
4310 TGTGAACCTGAACATCAACAGAGTGTCACTCAACCCGTGTAGAACACAGGCCACTGTGTG 4369
Qy
1400 AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db
4370 GATGAACATAAATCACTACAGTTGTAAATGTCCAGCAGGATTTTCAGGCCACAGGTGTGAG 4429
Qy
1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
Db
4430 ACAGACAGGCTTCCGGTTTAACTGGATTTTGAAGTTTCTGCACTACGGGTACGTC 4489
Qy
1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSer 1459
Db
4490 CTGCTAGATGGAGTGTGCCAACCTCCATGTCGTAACCTCGCATTTCTGGATGAAATCC 4549
Qy
1460 SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn 1479
Db
4550 TCTGATGTCACTAACATACGGGCGCCCATCTCTATGACATCTTGAGGATGACAAAGACAC 4609
Qy
1480 ThrLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIle 1499
Db
4610 ACCTTCTCTGCTGATTAACACGCTGGGTCTTTATGTGATGGAGGAAAGATC 4669
Qy
1500 ThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAlaIleThrThrThrSer 1519
Db
4670 ACCAATGCTCCCTCCGTAATATGATGTCATTTGGCATCATATTCATTCATTCAGCAAGT 4729
Qy
1520 AlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyValGlyLeu 1539
Db
4730 ATTTGGTGGAGCCCTGGAGGGTCTATATAGATGGGGAATTTATCTGACGGTGTACTGGCCCT 4789
Qy
1540 SerValGlyLeuProIleProGlyGlyValAlaLeuValLeuGlyGlnGluAspLys 1559
Db
4790 TCCATTGGCAAGCCATACCTTGTGGCGGTGCATTAGTTCTTGGCAAGACAGACAAA 4849
Qy
1560 LysGlyGlyGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579

4850 AAGGAGAGGGGTTCAACCCGGCTGAGTCTTTTGTGGCTCCATAAGCCAGCTCAACCTC 4909
Db
1580 TrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Qy
4910 TGGGACTATGTCCTGCTCCACAGCAGGTGAAGTTGCTGGCCAGCTCTGCCAGAGGA 4969
Db
1600 LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyValGlyLysVal 1619
Qy
4970 CTGAGTCGGGGAACGTTGACATGGCCCGCATTTCTCTGGGAATCAGGGGAGGTG 5029
Db
1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro 1639
Qy
5030 AAGGTTGATTCACGAGCATGTTCTGCTGATTCGCTCTTTAGAAAGATCCGCTGCT 5089
Db
1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
Qy
5090 CACCTGAGACCTGCAATCAGGAATCGAAAGCCAGGCTCCAAAGTCAGTCTGTCTGTAT 5149
Db
1660 ProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThr 1679
Qy
5150 CCGGCTTCAGATGTTGGGAATCCTGTGCAATGTTCTGACCAAGGGCAGTGGACA 5209
Db
1680 GlnProLeuProHisCysGluArgLysSerCysGlyValProProProLeuGluAsnGly 1699
Qy
5210 CAACCACTCCCACTGTAAGGCAATTCGCTGTGGCTGCTCCGCTTGGAGAAATGGC 5269
Db
1700 PheHisSerAlaAspAspPheThrAlaGlySerThrValThrTyrGlnCysAsnAsnGly 1719
Qy
5270 TTCTACTCAGCGGAGCTTCATCGGGGACACAGCGTGACCTATCAGTGCACCACTGGC 5329
Db
1720 TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyVal 1739
Qy
5330 TACTACTGCTGGGTGATTCGGAATGTTCTGCRAGACACAGGGAGCTGGACGGCATT 5389
Db
1740 SerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAla 1759
Qy
5390 TCACCATCTGCTGATGTCGATGATGTCAGTGGCTCGGCTCGGACTGTAGTAGCAGCACC 5449
Db
1760 SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp 1779
Qy
5450 TCCTGCTGTAACACACAGGATTCCTACGATGCTCTCTGTAAACCCACCATACACGGGAT 5509
Db
1780 GlyLysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHis 1799
Qy
5510 GGGAAATCTGCGAAGACCTGTAAATGTAAGGCTCCAGAAATCCAGAAATGGCCGC 5569
Db
1800 SerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGlyTyr 1819
Qy
5570 TCTTCTCGGAGATTACACCGTGGGTACTGTCAGTCACATTTCTGTGACGAGGAC 5629
Db
1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIle 1839
Qy
5630 GAGCTGGTGGGAGTGGACCATCATCGTGTGGGAGACTGGCGAGTGGGATCGCCTCAGG 5689
Db
1840 ProTyrCys 1842
Qy
5690 CGTCTGCT 5698
Db

RESULT 10

US-10-150-821-3
; Sequence 3, Application US/10150821
; Publication No. US20020192758A1
; GENERAL INFORMATION:
; APPLICANT: Welch, Andrew A.
; APPLICANT: Elliott, Gary S.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 01017/37592
; CURRENT APPLICATION NUMBER: US/10/150,821
; CURRENT FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: US/09/911,842
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438

; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 11230
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-150-821-3

Alignment Scores:

Pred. No.: 0 Length: 11230
Score: 8652.50 Matches: 1532
Percent Similarity: 90.56% Conservative: 137
Best Local Similarity: 83.13% Mismatches: 171
Query Match: 84.67% Indels: 3
Ds: 14 Gaps: 3

US-09-977-053-6 (1-1842) x US-10-150-821-3 (1-11230)

Qy 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
Db 176 ATGTGTCGGCTGGCTGCTTTTGTCTGGCTCTGGACTGTGGTGTGGGCTGGACCAAC 235
Qy 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
Db 236 TTCCAGCCCGTGGCCCTTGGCTCACTTCACTTCCGCTTCCGCTTCCGAGGCTCTCCG 295
Qy 41 GlyAlaProGlySerIleProAlaProProAlaProGly---AspGluAlaAlaGlySer 59
Db 296 GGGGCTCTGGGAGACTGGCGGTAGCTCCGCTCCAGTGGAGGAGGAGGAGGAGGAGC 355
Qy 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSer 79
Db 356 AAGTGGAGCGCTGGCGCGCGCTTCCGAGCCCGTGGCGAGCTGGCGAGCTCAGC 415
Qy 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPhe 99
Db 416 GCGAGCTTGGAGCTCGTCTTCTGCTGGACGAGCTGTCCAGCGTGGCGGCAACTTC 475
Qy 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValProThrAla 119
Db 476 CTCACGAGCTCAAGTTCGTGGCAAGCTGTCTCCGACTTCCCGTGGTGTCCAGCGCC 535
Qy 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
Db 536 ACGGTGTGGCAGTGTCACTTCTCATCCAGAACACAGCTGTGGTGGCGCGTGGATTAC 595
Qy 140 IleSerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluLeuProAla 159
Db 596 ATCTCCACGCGCGCGCACCAACACAGTGGCGCTGTCTCAGCGCGAGATCCCGGCG 655
Qy 160 IleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
Db 656 ATCACTTACCGCGGTGGTGGCAGCTATACAGGGCGCTTCCAGAACCGCGGCAATTC 715
Qy 180 LeuLeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSer 199
Db 716 CTTCCTGACTAGAGAAACTCCACCAAGTATATTTCTCATCACCGAGCTATTTC 775
Qy 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
Db 776 AATGGCGGAGACCCAGACCTATTCCAGCATCGCTTCCGGATTTCGGAGTGGAGATCTTC 835
Qy 220 ThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
Db 836 ACGTTCCGGGATTGGCAGGGGAAATCCGGAACTGAATGACATGGCTTCCACCCCGAAG 895
Qy 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
Db 896 GAAGAACAATGTACTGCTCCACAGTTTGAAGAATTTGAGGCTTTAGCTTCGAGGGCG 955
Qy 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
Db

956 TTGCATGAAGACTACCTTTCTGGGAGTATTCATCCAAAGAGGATATGCGCCACTGCTCTTAT 1015
Qy
280 LeuCysAspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHis 299
Db
1016 CTCGTGAGCGCTGGGAAGACTGCTGTGACAGAAATGGCCAGCTGCCAAATGTGGACACAC 1075
Qy
300 ThrGlyHisPheGluCysValCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGlu 319
Db
1076 ACGGTCATTTGATGATCTGTGAGAGGGCTATTACGGGAAGAGTCTGACAGCATGAG 1135
Qy
320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyLysSerSer 339
Db
1136 TGCACAGCTTGGCCATCAGGACATATAAGCCGGAAGCTTCTCCAGGAGGAATCAGCAC 1195
Qy
340 CysIleProCysProAspGluAenHisThrSerProGlySerThrSerProGluAsp 359
Db
1196 TGCATCCCATGCTCGACGTAAGCCACACTCCCTCCACCTGGGAAGCACTTCCCTGGAAGAC 1255
Qy
360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
Db
1256 TCGGTGTGCGAGAGGGATACAGAGATCTGGCCAGACCTGTGAGGTGTGTCCACTGTCT 1315
Qy
380 AlaLeuLysProProGluAenGlyTyrPheIleGlnAenThrCysAenAenHisPheAen 399
Db
1316 GCGCTGAAGCCTCTCAAAATGCTTTTATACAAAACACTTGCAAAACACTACTCTTCAAT 1375
Qy
400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu 419
Db
1376 GCGCGCTGTGGGTCCGATGTCCCGCGGCTTTGACCTTTGTGGGAAGCAGCATCCATTTG 1435
Qy
420 CysLeuProAenGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysPro 439
Db
1436 TGTCAACCCCAATGGTTGTGTCTGGACAGAAAGCTTCTGACAGATGAGAACTGCCCC 1495
Qy
440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys 459
Db
1496 CACCTCCGACAGCCCCAAACACAGCCCATCATCAGCTGCTCCACTCGGGAATGTCTTACAC 1555
Qy
460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr 479
Db
1556 ACCCTGTGTGTGTTACCTGCAATGAAGGATACAGATTAGAAGCGACACTAGGCTTACC 1615
Qy
480 CysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSer 499
Db
1616 TGTCAAGAAATGCCACAGTGGATGSCCCAGAGCCCGGTGTGTAGAACCCATTTGTGCC 1675
Qy
500 ThrPheGlnMetProLysAspValIleIleSerProHisAenCysGlyLysGlnProAla 519
Db
1676 ACCTTCCAGAGCCCCAAAGCGTTCATCTTCTCCACCCAGCTGCGGCAAGCAGCCGCC 1735
Qy
520 LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
Db
1736 AGGCGCTGGATGACCTGTCTAGCTAAGCTGCGCGCAGGATACATTTTATCCCGGGTCA 1795
Qy
540 GluMetLeuArgCysThrThrSerGlyLysThrPheAenValGlyValGlnAlaValCys 559
Db
1796 GAA--GTGAGTTGTGCCCATCTGGAGTGGAGTGCACCAAGTTCAGACAGCTGTGTGC 1852
Qy
560 LysAspValGluAlaProGlnIleAenCysPheLysAspIleGluAlaLysThrLeuGlu 579
Db
1853 AAGATGTGGAGCTCCACAAATCAGCTGTCCAAATGACATTGAGGCAAGACTGGGGAG 1912
Qy
580 GlnGlnAspSerAlaAenValThrTyrGlnIleProThrAlaLysAspAenSerGlyGlu 599
Db
1913 CAGCAGACTCTCTTAATGTCACCTGGCAAGTCCCAACAGCTAAAGCAACTCTGTGTAA 1972
Qy
600 LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
Db
1973 AAGGTGTGAGTGCAGCTGCACCCAGCTTTTACCCCACTTACCTTCTTCCCAATTTGGAGAC 2032
Qy
620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAenGlnAlaSerCysIlePheHis 639
Db
2033 GTGGCCATCACCTACAGCGCAACCGACTCATCTCCGGTAAACCAAGCCAGCTGCACTTCTAC 2092

640 IleLysValIleAspAlaGluProProValIleAspTyrCysArgSerProProProVal 659
Db
2093 ATTAAGGTCAATGATGTGGAAACCGCTGTCTATAGATTGGTCCGATCTCCACTCCCAATC 2152
Qy
660 GlnValSerGluLysValHisAlaSerTyrAspGluProGlnPheSerAspAenSer 679
Db
2153 CAGGTCTGAGAGAAGCAGCACCCCTGCAAGCTGGATGAGCTCAGTTCTCAGACAACTCC 2212
Qy
680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
Db
2213 GGGGCTGAAATGTGCTCAATACACAGCTGACACAAAGGCGACATGTTCTCATGGGAA 2272
Qy
700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAenAenArgThrCysAspIleHis 719
Db
2273 ACGGTGGTGTGTACACAGCACTGACCCCTCAGGCNAACAGGAGCTGTGACATCCAC 2332
Qy
720 IleValIleLysGlySerProCysGluIleProPheThrProValAenGlyAspPheIle 739
Db
2333 ATTGTCATAAAGGTTCTCCCTGTGAGTCCCTTCAACCCCTGTATAACGGGAGCTTTATC 2392
Qy
740 CysThrProAspAenThrGlyValAenCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
Db
2393 TGTGCCCGAGATAGTGTGGAGTTAACTGTAGCTGTGAGCTGCAAGGAGGCTATGATTC 2452
Qy
760 ThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTyrLysProThr 779
Db
2453 ACGAAGGGTCACTGAGAAGTACTACTGTGCTTTTGNAGATGTATCTGGAGACCA 2512
Qy
780 TyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAenHisGlyPheLysSer 799
Db
2513 TACTCTACAGATGGCCAGACTGTGCTATAAAGCTTTTGCACCAACCATGTTTCAAGTCC 2572
Qy
800 PheGluMetPheTyrLysAlaAlaAaCysAspAspThrAspLeuMetLysLysPheSer 819
Db
2573 TTTGAAATGCTATACAAAACCACTCGCTGTGTGATGACATGGAATCTGTTTAAAGAGTTTCT 2632
Qy
820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
Db
2633 CGAGCATTTGAGACTACCTCGGGGAACATGCTCCGCTCTTTGTACGATGCTGATGAC 2692
Qy
840 IleAspCysArgLeuGluAenLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
Db
2693 ATTGACTGCAGACTGGAGGAC---CTGACCAAAAATACTGCACTGAGTATAATTAACAC 2749
Qy
860 TyrGluAenGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAenAaGLeuAspTyr 879
Db
2750 TATGAAATGGCTTTGCAATTGGACCAAGGAGCTGGGTGAGCAACAGGCTGATTTAT 2809
Qy
880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
Db
2810 TCTTACATCACTTCTCGATGTTGTACAGAAACACCCACCGATGTGGCAAGCCAGA 2869
Qy
900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAen 919
Db
2870 TCGTCAAGGATTAAGAAGCACTGCTCCCATTTCTGACCCCAAAATTCAGCTAATTTTAAAC 2929
Qy
920 IleThrAlaSerValProLeuProAspGluArgAenAspThrLeuGluTyrGluAenGln 939
Db
2930 ATCACAGTACGCTGCCACTCCAGAGGAAGAAACGATACCTTGAATTCGAGAACTCAG 2989
Qy
940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAenLys 959
Db
2990 CAGGACTCATTAAGACATTGGAAACAACTACCAATCGCTGAAAAGCACCCTTGNATAA 3049
Qy
960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
Db
3050 GAGCCCATGTATTCTTTCCAGCTGCGCTCGAAGACAGTGGGTGGCTGACAGCAATTCCTC 3109
Qy
980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
Db
3110 GAAACAGAAAGGCTTTTCTCTTCTGACAGCAGGCTCTGTGTGCTGAGGGGCGCATGTGT 3169

QY 1000 ValAenCysProLeuGlyThrTyrrAenLeuGluHisPheThrCysGluSerCysArg 1019
Db 3170 GTCACCTGCCCCCTGGGAACCTCTTACTCTCTGGAGCAATCCACCTGTGAAGCTGCGCTC 3229
QY 1020 IleGlySerTyrrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
Db 3230 ATGGGATCTACCAAGATGAAGAGGCGAGCTGGAATGCAAGCTCTGTCCCCCAAGGACT 3289
QY 1040 TyrThrGluTyrrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysIysGlnGly 1059
Db 3290 CACGGGAATACCTCCATTCAGGAAGCGTCTCTGAATGCAAGCTCAGTGTGAAGCAAGGC 3349
QY 1060 ThrTyrrSerTyrrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrrGlnPro 1079
Db 3350 ACCTACTCTTCCAGTGGGCTGGAGACCTGCGAATCGTGTCCGCTGGGTACTTATCAACCG 3409
QY 1080 LysPheGlySerArgSerCysLeuSerCysProGluAenThrSerThrValLysArgGly 1099
Db 3410 GAAATTTGGATCCCGGAGCTGCTCTATGCCAGAAACCCACCAACCGTGAAGAGGCA 3469
QY 1100 AlaValAenIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
Db 3470 GCCGTGGACATCTCTGCTGTGGAGTGGCTGCCAGTAGGAGAAATCTCCCGTCTCTGG 3529
QY 1120 LeuMetProCysHisProCysProArgAspTyrrTyrrGlnProAsnAlaGlyLysAlaPhe 1139
Db 3530 CTAACACCTGCTACCTCTGACGACTATACCAACCCCAATGCGAGGAGAGTCTCTTC 3589
QY 1140 CysLeuAlaCysProPheTyrrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
Db 3590 TGCCTGGCTGTGCTCTTTATGGAATACCAACCATCATCTGGCGCCACCGTCCATCACAGC 3649
QY 1160 CysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValProAlaSer 1179
Db 3650 TGCTCAAGTTTTAGCTCTACTTTCTCAGCAGCAGAGAAGACATAGTCCCTCGTGCC 3709
QY 1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe 1199
Db 3710 CCTGGACATTCCTCAGAACAGTACGAGTACAGCTCAGGCTTTTCAGCAATGCTCTCTTA 3769
QY 1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrrValCysLeuCys 1219
Db 3770 AACCCCTGCCCAACACAGTGGAACTGCCACACAGCTTGGCGGTGATGTCGTCTCTGCG 3829
QY 1220 ProLeuGlyTyrrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
Db 3830 CCACCTGGATACACAGGCTTAAAGTGTGAAGTGAATATGATGATGACAGCTCTCTGCT 3889
QY 1240 CysLeuAenAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
Db 3890 TGCCTCAATGTGTGAATTTGTAGAACCCAGTTGGGGGATTCACGTGGCAATGTTTCATG 3949
QY 1260 GlyTyrrThrGlyGlnArgCysGluGluAenIleAsnGluCysSerSerProCysLeu 1279
Db 3950 GGCTATTGAGGTCAAAATGATGCAAGAAATATAATGATGATGATCTCCAGGCCCTTGCTTA 4009
QY 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrrArgCysThrCysValLysGlyPhe 1299
Db 4010 AATAAAGGAACCTGCATGACGGCTTGGCAAGCTACCGCTGTACCTGTGTGAAGAGATAC 4069
QY 1300 ValGlyLeuHisCysGluThrGluValAenGluCysGlnSerAsnProCysLeuAenAsn 1319
Db 4070 ATGGGTGTGCACTGTGAACAGACGTCAATGATGATGATGATGATGATGATGATGATGAT 4129
QY 1320 AlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGly 1339
Db 4130 GCAGTTTGTAAAGACCAAGTTGGGGGTCTCGTGCATAATGCGCAATGCCAGTCCCTGCTTAACAC 4189
QY 1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr 1359
Db 4190 ACTCGGTGTGAAGAAATGTGGATGAGTGTCTCAGTCAGCCATGCCAAATATGGAGCCACT 4249
QY 1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379

Db 4250 TGTAAAGATGTTGCCAACAGCTTCAGGTGTCAATGTCCAGCAGGCTTCACAGGGACACAC 4309
QY 1380 CysGluLeuAenIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
Db 4310 TGTGAACCTGAACANTCAACGAGTGTGCTCAACCCGCTGTAGAACCCAGGCCACTGTGTG 4369
QY 1400 AspGluLeuAenSerTyrrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db 4370 GATGAACCTAAACTCATACAGTTGTAAATGTACGACGAGATTTTCAGGCCACAGGTGTGAG 4429
QY 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrrGlyTyrrVal 1439
Db 4430 ACAGAACAGCTCTCCGGTTTAACTCGATTTTGAATTTCTGGCATCTACGGGTACGTC 4489
QY 1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSer 1459
Db 4490 CTGCTAGATGGAGTGTCTGCCAACCTCCATGCGTAAACCTGGGCATTTCTGGATGAATCC 4549
QY 1460 SerAspAspMetAsnTyrrGlyThrProIleSerTyrrAlaValAspAsnGlySerAspAsn 1479
Db 4550 TCTGATGTCATCACTACGAGACGCCATCTCTCTATGCACTTGAGGATGCAACAGACAC 4609
QY 1480 ThrLeuLeuLeuThrAspTyrrAsnGlyTrpValLeuTyrrValAenGlyArgGluLysIle 1499
Db 4610 ACCTTCCTCTCTGACTGATTTACAAACGGCTGGGTCTTTATGTGAATGGAAGGAAAGATC 4669
QY 1500 ThrAenCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSer 1519
Db 4670 ACCAACTGCCCTCCGTAATGATGGCATTTTGGCATCATATTTGCAATCATCATGGACAGT 4729
QY 1520 AlaAenGlyIleTrpLysValTyrrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeu 1539
Db 4730 ATTGTTGAGGCTGAGGGTCTATATAGTGGGAAATATCTGACGGTGTACTGCGCTC 4789
QY 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLys 1559
Db 4790 TCCATTGGCAAGGCCATACCTGGTGGGTGCTAGTTCTTTGGGCAAGAGCAAGACAA 4849
QY 1560 LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAenLeu 1579
Db 4850 AAAGGAGAGGGGTTCACACCGGCTGAGTCTTTGTGGGTGCTCATTAAGCAGCTCAACCTC 4909
QY 1580 TrpAspTyrrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4910 TGGGACTATGTCCTGTCTCCACAGCAGGAGTGAAGTGTCTGGCCAGCTCTCTCCAGAGGAA 4969
QY 1600 LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysVal 1619
Db 4970 CTGAGTCGGGGAACGTTGTAGCATGGCCGATTTCTGTCCGGAATCAAGGGGAAGGTG 5029
QY 1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyCysValPro 1639
Db 5030 AAGTTTGAATTCAGCAGCATGTTCTGTCTGATTTGTCCGCTCTTTAGAACGATCCGTGCT 5089
QY 1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
Db 5090 CACCTGAGACCTGTCATCAGAAATCGAAAGCCAGCTCCAAAGTCAGTCTGTTCTGTGAT 5149
QY 1660 ProGlyPheGlnLeuValGlyAenProValGlnTyrrCysLeuAenGlnGlyGlnTrpThr 1679
Db 5150 CCGGCTTCCAGATGGTGGGAATCTCTGTCAGTATTTGTCTGAACCAAGCGGCGTGGACA 5209
QY 1680 GlnProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAenGly 1699
Db 5210 CAACCACTCCCCACTGTGAACGCAATCGCTGTGGGCTGCCCTCCGCCCTTGGAGATGGC 5269
QY 1700 PheHisSerAlaAspAspPheTyrrAlaGlySerThrValThrTyrrGlnCysAsnAsnGly 1719
Db 5270 TTTACTCAGCGAGACTTCCATCGGGGACAGCAGGTGACCTATCAGTGCACCACTGGC 5329
QY 1720 TyrTyrrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAenGlyVal 1739

```
Db 5330 TACTACCTGCTGGTGAATCCCGAATGTTCTGCRKAGACAAACCGAGCTGGAACGGCATT 5389
Qy 1740 SerProSerCysLeuAspValAspGluCysAlaValAlaGlySerAspCysSerGluHisAla 1759
Db 5390 TCACCATCTCTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 5449
Qy 1760 SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp 1779
Db 5450 TCTGCTGCTGACACCAACGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 5509
Qy 1780 GlyLysAsnCysAlaGluProLysCysLysAlaProGlyAsnProGluAsnGlyHis 1799
Db 5510 GGGAAAACTGTCGACAACTGTAAATGTAAGGCTCCAGAAAATCCAGAAAATGCGCGC 5569
Qy 1800 SerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyr 1819
Db 5570 TCTTCTGCGCAGATTACACCGCTGGGTACTGTCAGTGCATGTCATGTCATGTCATGTCATGTCATGTCAT 5629
Qy 1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluThrPheAsnHisLeuIle 1839
Db 5630 GAGCTGCTGGAGTGAGCACCATCACGTGTTGGAGACTGGCGAGTGGGATCCCTCAGG 5689
Qy 1840 ProTyrCys 1842
Db 5690 CCGTCTCTGT 5698

RESULT 11
US-10-302-172-34
; Sequence 34, Application US/10302172
; Publication NO. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCp
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (415)..(2991)
US-10-302-172-34

Alignment Scores:
Pred. No.: 0 Length: 3448
Score: 5485.00 Matches: 982
Percent Similarity: 99.90% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 53.67% Indels: 0
Db: 13 Gaps: 0

US-09-977-053-6 (1-1842) x US-10-302-172-34 (1-3448)
Qy 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIle 337
Db 1 TATGAATGACAGCTGTCCTCCATCGGGGACATCAAACTGGAAGCTCCACAGAGAAATC 60
Qy 338 SerSerCysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerPro 357
Db 61 AGCAGTTGCATTCATTCCTCCGATGTCCTGAAATCACACCTCTCCACCTTGGAGGACATCCCT 120
```

```
Qy 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
Db 121 GAAGACTGTGTCTGTCAGAGAGGGATACAGGGCATCTGGCCAGACCTGTGAATCTTGTCCAC 180
Qy 378 CysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnHis 397
Db 181 TGGCCTGCCCTGAAGCTCCCGAAATGGTTACTTTATCCAAACACTTGGCAACACCAC 240
Qy 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIle 417
Db 241 TTCAATGACAGCTGTGGGTCCGATGTCACCTGGAATTTGATCTTGTGGGAAGCAGCATC 300
Qy 418 IleLeuCysLeuProAsnGlyLeuThrSerGlySerGluSerTyrCysArgValArgThr 437
Db 301 ATCTTATGTCATCCCAATGGTTTGTGGTCCGTTTCAGAGAGCTACTGCAGAGTAAGAACA 360
Qy 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
Db 361 TGTCTCTCATCTCCGACCGCAACATGGCCACATCAGCTGTCTACAGGGGAAATGTTA 420
Qy 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
Db 421 TATAAGACAACATGTTTGGTTCCTGTGATGAAGGGTACAGACTAGAAGGCGAGTATAAG 480
Qy 478 LeuThrCysGlnGlyAsnSerGlnThrAspGlyProGluProArgCysValGluArgHis 497
Db 481 CTTTACTTGTCAAGGAACACAGCTGGATGGGCGCAGACCCCGGTGTGTGGAGCGCCAC 540
Qy 498 CysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGln 517
Db 541 TGTTCACACTTTCAGATGCCAAGATGTCTATCATATATATATATATATATATATATATAT 600
Qy 518 ProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGly 537
Db 601 CCAGCCAAATTTGGGACGATCTCTATGTAAGTTGCGCCCAAGGGTTCATTTTATCTGGA 660
Qy 538 ValLysGluMetLeuArgCysThrThrSerGlyThrPheAsnValGlyValGluAlaAla 557
Db 661 GTCAAGAAATGCTGAGATGTACCACTTCTGGAATAATGGAAATGTGGAGTTTCAGGAGCT 720
Qy 558 ValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThr 577
Db 721 GTGTGTAAGAGCGTGGAGCTCTCTCAAAATCAACTGTCTTAAGGACATAGAGGCTAAGACT 780
Qy 578 LeuGluGlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSer 597
Db 781 CTGGAACAGCAAGATTTCTGCAATGTTTACCTGGCAGATTCCAACAGCTAAAGACAACCTCT 840
Qy 598 GlyGluLysValSerValHisValHisProAlaPheThrProTyrLeuPheProIle 617
Db 841 GGTGAAGAGGTGTCACTGTCAGTTCACCTGTTTCACTCCACCTTTTACCTTTTCCCAATT 900
Qy 618 GlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
Db 901 GGAGATGTTGCTATCGTATACAGGCACTGACTATCCGGCAACCGGCGAGCTGCTGATT 960
Qy 638 PheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProPro 657
Db 961 TTCCCATATCAAGGTTATTGATGTCAGAAACCACTCTGTCTATAGACTGGTGCAGATCTCCACCT 1020
Qy 658 ProValGlnValSerGluLysValHisAlaAlaSerThrPheAspGluProGlnPheSerAsp 677
Db 1021 CCCGTCCAGGTCTCGGAGAGGTACATGCGCAGCTGGAGTGGGATGAGCTCTGTTCTCAGAC 1080
Qy 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697
Db 1081 AACTCAGGGGCTGAAATGTTGTCATTACAGAGTCTATACACAGAGACCTTTTCCCTCAA 1140
Qy 698 GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717
Db 1141 GGGGAGACTATAGTACATATACAGCACTGACCCCTCAGGCAATTAACAGGACATGTGAT 1200
```

QY 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
DB 1201 ATCCATATGTGTCATAAAAGGTTCTCCCTGTGAAATTCACATTCACACCTGTAAATGGGGAT 1260
QY 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
DB 1261 TTTATATGCATCCAGATTAATCTGGAGTCACTGACATTAACCTGCTTGGAGGGCTAT 1320
QY 758 AspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLys 777
DB 1321 GATTTTCACAGAAGGCTTACTGACAGATTAATTTGCTTATGAAGATGGCGTCTCGAAA 1380
QY 778 ProThrTyrThrGluTrpProAspCysAlaLysLysAspPheAlaAsnHisGlyPhe 797
DB 1381 CCAACATATACCACTGAATGGCCAGACTGTGCCAAAAACGTTTTCGAAACCAACGGGTT 1440
QY 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspThrAspLeuMetLysLys 817
DB 1441 AAGTCTTTGAGATGTTCTACAAAGCAGCTGCTGTGATGACACAGATCTGATGAAGAAG 1500
QY 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
DB 1501 TTTTCTGAAGCATTTTCAGACGACCTGGGAAATGGTCCCATCATTTTGTAGTGATGCA 1560
QY 838 GluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsn 857
DB 1561 GAGGACATGTGACTGACAGCTGGAGGAGAACCTGCACAAAATAATTTGCCTAGAATATAT 1620
QY 858 TyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAlaAsnArgLeu 877
DB 1621 TATGACTATGAATAATGGCTTTGCAATGGACAGGTGGCTGGGGTGCAGCTAATAGGCTG 1680
QY 878 AspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsn 897
DB 1681 GATTAATCTTACAGTACTCTCTGACACTGTCGCAAGTGTGCAAGAAACAGCCACAGCATCGSCAAT 1740
QY 898 AlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIle 917
DB 1741 GCCAGGTCTCTACAGGATTAAGAAGAGTGGCCCATTAATCTGACTATAAATAATTAAGTTAAT 1800
QY 918 PheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu 937
DB 1801 TTTAACATCACAGCTAGTGTGCTTACCCGATGAAGAAATGATACCTCTGAATGGGAA 1860
QY 938 AsnGlnGlnArgLeuLeuGlnThrIleThrAsnLysLeuLysArgThrLeu 957
DB 1861 AATCAGCAACCACTCTCTCAGACATTTGGAATCTATCACAAATAAATCTGAAAAGGCTCTC 1920
QY 958 AsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsn 977
DB 1921 AACAAAGACCCCAATGATTTCTCTTTCAGCTTGCTCAGAAATACCTATAGCCACAGCAAT 1980
QY 978 SerLeuGluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArg 997
DB 1981 TCATTAGAAACAAAAAGGCTTCCCTCTCTCAGACACAGGCTCAGTGTGAGAGGGCT 2040
QY 998 MetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer 1017
DB 2041 ATGTGTGTCATTTGGCCCTTTGGGAACCTATTATATATCTGGAACATTTACCTGTGAAAGC 2100
QY 1018 CysArgIleGlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSer 1037
DB 2101 TGCCGGATCGGATCTCTATCAGATGAGAGGGCACTTGAATGACAGCTTGGCCCTCT 2160
QY 1038 GlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLys 1057
DB 2161 GGGATGTACAGGAATATATCATTTCAAGAAACATCTCTGATTTGTAAGCTCAGTGTAAA 2220
QY 1058 GlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyr 1077
DB 2221 CAAGGCACCTACTCATACAGTGGACTTGGAGCTTGTGAATCTGTGCTCCACTGGGCACTTAT 2280
QY 1078 GlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLys 1097

DB 2281 CAGCCCAAAATTTGGTTCCGGAGCTGCCTCTCTGTCTCAGAAAACACCTCACTGTGAAA 2340
QY 1098 ArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArg 1117
DB 2341 AGAGGAGCGGTGAACATTTCTGCATGTGGAGTTCTTGTCCAGAGGAAAATTTCTCGCT 2400
QY 1118 SerGlyLeuMetProCysHisProCysProAspTyrTyrGlnProAsnAlaGlyLys 1137
DB 2401 TCTGGGTTAATGCCCTGTCCACCATGTCTCTGACTATTACCACTAATGACGGAAG 2460
QY 1138 AlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIle 1157
DB 2461 GCTTCTGCTGGCTGCTGCTTATATGAACTACCCCAATTCCTGCTGCTCCAGATCCATC 2520
QY 1158 ThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluSerValValProPro 1177
DB 2521 ACAGAATGTTCAAGTTTATGTTCACTTCTCAGCGGAGGAAAAGTGTGTGCCCT 2580
QY 1178 AlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCys 1197
DB 2581 GCCTCTCTTGGACATTAATAAAGAGGCATGAATCAGCAGTCAGGTTTCCATGATATGC 2640
QY 1198 PhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCys 1217
DB 2641 TTTCTTAACCTTTGCCACATAGTAGTGENAACCCTGCCAGCAACTTGGGCTGTATGTTGT 2700
QY 1218 LeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerPro 1237
DB 2701 CTCTGTCCACTTGGATATACAGGCTTAAAGTGTGAAACAGACATCATGATGAGTGCAGCCA 2760
QY 1238 LeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCys 1257
DB 2761 CTGCTTCTGCTCAACATGAGGTTTGTAAAGACCTAGTTGGGGAATTCATTGTGAGTGC 2820
QY 1258 ProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerPro 1277
DB 2821 CCATCAGGTTACACAGGTCAGCGGTGTGAAGAAAATATAAATGAGTGTAGTCCAGTCT 2880
QY 1278 CysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLys 1297
DB 2881 TGTTTTAAATAAAGGAATCTGTGTGATGTTGTGCTGGCTGCTATCGTTGCCACATGTGTGAAA 2940
QY 1298 GlyPheValGly 1301
DB 2941 GGAATTGTAGT 2952

RESULT 12

US-09-822-846-55/c
; Sequence 55, Application US/09822846
; Publication No. US20030027139A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fectel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400


```

; APPLICANT: Agostino, Michael J.
; APPLICANT: Steininger II, Robert J.
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Pechtel, Kim
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6400
; CURRENT APPLICATION NUMBER: US/09/822,846
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,605
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 629
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 1408
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-846-54

Alignment Scores:
Pred. No.: 1,21e-189 Length: 1408
Score: 1977.00 Matches: 362
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 1
Query Match: 19.35% Indels: 1
DB: 10 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-822-846-54 (1-1408)

Qy 505 LysAspValIleLeuSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThrIle 524
Db 2 AARGATGTCATCATATCCCCCACAACTGTGGCAGCAGCCAGCCAAATTTGGGACGATC 61

Qy 525 CysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArgCys 544
Db 62 TGTATGTAAATGTCGGCCAGGGTTTCATTTTATCTGGAGTCAAGAAATGCTGAGATGT 121

Qy 545 ThrThrSerGlyLysTyrPasnValGlyValGlnAlaLysValCysLysAspValGluAla 564
Db 122 ACCACTTCTGGAATGGAATGTCGGAGTTTCAGGCAGCTGTGTGTAAGACGTGGAGGCT 181

Qy 565 ProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnGlnAspSerAla 584
Db 182 CCTCAATCAACTGTCCTTAAGGACATAGAGGCTAGAGCTCTGGACAGCAAGATTCGCG 241

Qy 585 AsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerValHis 604
Db 242 AATGTTACTTGGCAGATTCACAGCTAAAGACAACTCTGGTGAAGAGGTGTCAGTCCAC 301

Qy 605 ValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleValTyr 624
Db 302 GTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATTTGGAGATGTTGCTATCGTATAC 361

Qy 625 ThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIleAsp 644
Db 362 ACGGCAACTGACCTATCCGGCAACACAGGCCAGCTGCTATTTTCATATCAGGATTTATAT 421

Qy 645 AlaGluProProValIleAspTyrCysArgSerProProProValGlnValSerGluLys 664
Db 422 GCAGAACCACTGTCATAGACTGGTGACATCTCCACCTCCCGTCCAGGCTCGGAGAG 481

Qy 665 ValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeuVal 684
Db 482 GTACATGCCGCAAGCTGGGATGAGCTCAGTCTCTCAGACAACTCAGGCGCT-GAATTTGGTC 540

Qy 685 IleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGlnTyr 704

```

```

Db 541 ATTACCAAGAGTCATACACAAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAGTAT 600
Qy 705 ThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGly 724
Db 601 ACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATTTGTCTAAGAAGT 660
Qy 725 SerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAspAsn 744
Db 661 TCTCCCTGTGAATTCATTCACACCTGTAATGAGGGATTTTATATGCACTCCAGATAT 720
Qy 745 ThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySerThr 764
Db 721 ACTGGAGTCAACTGTACATTAACCTTGTGGAGGGCTATGATTTTCACAGAGGGTCTACT 780
Qy 765 AspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrThrGluTrp 784
Db 781 GACCAAGTATTATTGTGCTTATGAAGATGGCGTCTGGAAACCAACATATATACCACTGAATGG 840
Qy 785 ProAspCysAlaLysLysAspGlyPheAlaAsnHisGlyPheLysSerPheGluMetPheTyr 804
Db 841 CCAGACTGTGCCAAAAACGGTTTTCGAAACCCACGGGTTCAAGTCTTTGAGATGTTCTAC 900
Qy 805 LysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGluThr 824
Db 901 AAGCAGACTCGTTGTGATGACACAGATCTGTATGAAGAGTTTCTGAAGCAATTTGAGACG 960
Qy 825 ThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArgLeu 844
Db 961 ACCCTGGGAAAAATGTCCTCCATCATTTTGTAGTATGACAGAGACATTCACCTGCAGACTG 1020
Qy 845 GluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe 864
Db 1021 GAGGAGAACCTTCGACCAAAATATTTGCTAGAAATATATATATGACTATGAAATGGCTTT 1080
Qy 865 AlaIleGly 867
Db 1081 GCATTTGGT 1089

RESULT 14
US-09-764-898-49
; Sequence 49, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-898-49

Alignment Scores:
Pred. No.: 6,26e-141 Length: 1892
Score: 1498.00 Matches: 263
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 14.66% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-764-898-49 (1-1892)

Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
Db 2 TGTGTGTATGGGTGGCTGGCTGCTATCTGTCACATGTGTGAAGGATTTGTAGCGCTGCAT 61

Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu 1323

```



```

Alignment Scores:
Pred. No.: 2.4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 0
Query Match: 12.51% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-764-898-122 (1-680)

QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
DB 3 TGTGTTGATGGTGGCTGCTATCTTCACATGTCGTAAGGATTTGTAGCCCTGCAT 62

QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
DB 63 TGTGAAACAGAGTCAATGAATGCCAGTCAACCCATGCTTAATAATATGCGTCTGTGAA 122

QY 1324 AspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
DB 123 GACCAAGTTCGATGAGTCTCATGTCAGCCATGCAAAATGAGAGTACCTGTAAAGACGGT 182

QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
DB 183 AAGAACGTTCGATGAGTCTCATGTCAGCCATGCAAAATGAGAGTACCTGTAAAGACGGT 242

QY 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
DB 243 GCCAATAGCTTCAGATGCTGTGTGCGAGCTGGCTTCACAGGATCACACTGTGAATGAAC 302

QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
DB 303 ATCAATGATGTGAGTCTCAATCCATGTAGAAATCAGGCCACCTGTGTGAATGAAT 362

QY 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlnSer 1423
DB 363 TCATACAGTTGTAAATGTGAGCAGGATTTTCAGGCMAAAGGTGTGAAACAGACAGTCT 422

QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
DB 423 ACAGGCTTTTAACTCGGATTTTGAGTTTCGGCATCTATGANTATGTCTAGTAGTGC 482

QY 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
DB 543 AACTATGGAAACCAATCTCTATGAGTGTGATACGGCAGCGACATACCTTGTCTCTG 602

QY 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
DB 603 ACTGATTATAACGGCTGGGTTCCTTATGTAATGGCAGGAAAGATAACAACACTGTCCC 662

QY 1504 SerValAsnAspGlyArg 1509
DB 663 TCGGTGAATGATGGCAGA 680

RESULT 17
US-09-764-881-14
; Sequence 14, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14

Alignment Scores:
Pred. No.: 2.4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 0
Query Match: 12.51% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-764-881-14 (1-680)

QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
DB 3 TGTGTTGATGGTGGCTGCTATCTTCACATGTCGTAAGGATTTGTAGCCCTGCAT 62

QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
DB 63 TGTGAAACAGAGTCAATGAATGCCAGTCAACCCATGCTTAATAATATGCGTCTGTGAA 122

QY 1324 AspGlnValGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
DB 123 GACCAAGTTCGATGAGTCTCATGTCAGCCATGCAAAATGAGAGTACCTGTAAAGACGGT 182

QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
DB 183 AAGAACGTTCGATGAGTCTCATGTCAGCCATGCAAAATGAGAGTACCTGTAAAGACGGT 242

QY 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
DB 243 GCCAATAGCTTCAGATGCTGTGTGCGAGCTGGCTTCACAGGATCACACTGTGAATGAAC 302

QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
DB 303 ATCAATGATGTGAGTCTCAATCCATGTAGAAATCAGGCCACCTGTGTGAATGAAT 362

QY 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlnSer 1423
DB 363 TCATACAGTTGTAAATGTGAGCAGGATTTTCAGGCMAAAGGTGTGAAACAGACAGTCT 422

QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
DB 423 ACAGGCTTTTAACTCGGATTTTGAGTTTCGGCATCTATGANTATGTCTAGTAGTGC 482

QY 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerAspMet 1463
DB 483 ATGCTCCCATCTCTCCATGCTTAACCTGTACCTTCTGGATGAAATCCCTGACGACATG 542

QY 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
DB 543 AACTATGGAAACCAATCTCTATGAGTGTGATACGGCAGCGACATACCTTGTCTCTG 602

QY 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
DB 603 ACTGATTATAACGGCTGGGTTCCTTATGTAATGGCAGGAAAGATAACAACACTGTCCC 662

QY 1504 SerValAsnAspGlyArg 1509
DB 663 TCGGTGAATGATGGCAGA 680

RESULT 18
US-09-764-881-14
; Sequence 14, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14

```


; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 192

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 680

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-881-14

Alignment Scores:

Pred. No.: 2,4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 12.51% Indels: 0
DB: 13 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-764-881-14 (1-680)

```
QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
DB 3 TGTGTTGATGGTGGCTGCTATCGTTGCACATGTGTGAAGGATTTGTAGGCCCTGCAT 62

QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
DB 63 TGTGAACACAGAGTCATGATGCCAGTCACCCATGCTTAATATGATGAGTCTGTGAA 122

QY 1324 AspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
DB 123 GACCAAGTTGGGGGATTCATGTGCAATGCCAGTCACCCATGCTTAATATGATGAGTCTGTGAA 182

QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
DB 183 AAGAACGTCGATGAGTCTCAGTCAGCCATGCCAATGCGAGTACCTGTAAAGACGGT 242

QY 1364 AlaAsnSerPheArgCysLeuCysAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
DB 243 GCCAATAGCTTCAGATGCTGTCAGTCAGCCATGCCAATGCGAGTACCTGTAAAGACGGT 302

QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
DB 303 ATCAATGATGTGAGTCTTAATCCATGTAGAAATCAGGCCACCTGTGTGATGAATTAAT 362

QY 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyAlaArgCysGluThrGlnSer 1423
DB 363 TCATACAGTTGTAAATGTGACCCAGGATTTTCAGGCGAAAGGTGTGAACAGACAGTCT 422

QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
DB 423 ACAGGCTTTAACTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCTAGATGGC 482

QY 1444 MetLeuProSerLeuHisAlaLeuThrCysThrTrpMetLysSerSerAspMet 1463
DB 483 ATGCTCCCATCTCTCCATGCTCTAACCTGTACCTCTGAGTGAATCTCTGACGACATG 542

QY 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
DB 543 AACTATGGAACCAATCTCTATGACAGTTGATAACGGCAGCGACATACCTTGTCTCTG 602

QY 1484 ThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
DB 603 ACTGATTATACGGCTGGGTTCTTTATGTGATGCGCGGGAAGATATACAACTGTCTCC 662

QY 1504 SerValAsnAspGlyArg 1509
DB 663 TCGGTGAATGATGGCAGA 680
```

RESULT 19

US-09-764-893-37

; Sequence 37, Application US/09764893

; Publication No. US20020086330A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PJ209

; CURRENT APPLICATION NUMBER: US/09/764,893

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 154

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 37

; LENGTH: 680

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-893-37

Alignment Scores:
Pred. No.: 2,4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 12.51% Indels: 0
DB: 13 Gaps: 0

US-09-977-053-6 (1-1842) x US-09-764-893-37 (1-680)

```
QY 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
DB 3 TGTGTTGATGGTGGCTGCTATCGTTGCACATGTGTGAAGGATTTGTAGGCCCTGCAT 62

QY 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
DB 63 TGTGAACACAGAGTCATGATGCCAGTCACCCATGCTTAATATGATGAGTCTGTGAA 122

QY 1324 AspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
DB 123 GACCAAGTTGGGGGATTCATGTGCAATGCCAGTCACCCATGCTTAATATGATGAGTCTGTGAA 182

QY 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
DB 183 AAGAACGTCGATGAGTCTCAGTCAGCCATGCCAATGCGAGTACCTGTAAAGACGGT 242

QY 1364 AlaAsnSerPheArgCysLeuCysAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
DB 243 GCCAATAGCTTCAGATGCTGTCAGTCAGCCATGCCAATGCGAGTACCTGTAAAGACGGT 302

QY 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
DB 303 ATCAATGATGTGAGTCTTAATCCATGTAGAAATCAGGCCACCTGTGTGATGAATTAAT 362

QY 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyIleTyrGlyTyrValMetLeuAspGly 1423
DB 363 TCATACAGTTGTAAATGTGACCCAGGATTTTCAGGCGAAAGGTGTGAACAGACAGTCT 422

QY 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
DB 423 ACAGGCTTTAACTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCTAGATGGC 482

QY 1444 MetLeuProSerLeuHisAlaLeuThrCysThrTrpMetLysSerSerAspMet 1463
DB 483 ATGCTCCCATCTCTCCATGCTCTAACCTGTACCTCTGAGTGAATCTCTGACGACATG 542

QY 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
DB 543 AACTATGGAACCAATCTCTATGACAGTTGATAACGGCAGCGACATACCTTGTCTCTG 602

QY 1484 ThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
DB 603 ACTGATTATACGGCTGGGTTCTTTATGTGATGCGCGGGAAGATATACAACTGTCTCC 662

QY 1504 SerValAsnAspGlyArg 1509
DB 663 TCGGTGAATGATGGCAGA 680
```

RESULT 20

US-10-073-865-37

; Sequence 37, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: RJ209C1
; CURRENT APPLICATION NUMBER: US/10/073,865
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 37
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-073-865-37

Alignment Scores:
Pred. No.: 2,4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 12.51% Indels: 0
DB: 15 Gaps: 0

US-09-977-053-6 (1-1842) x US-10-073-865-37 (1-680)

QY	1284	CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis	1303
DB	3	TGTGTTGATGTTGGCTATGTTGCACATGTTGAAGAGGATTTGTAGCCCTGCAT	62
QY	1304	CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu	1323
DB	63	TGTGAACAGAGAGTCAATGATGCAATGCTTAAATATGCACTGTGTGAA	122
QY	1324	AspGlnValGlyPheLeuCysValProGlyPheLeuGlyThrArgCysGly	1343
DB	123	GACCAAGTTGGGGATTCATGTGCAATGCCACCTGGATTTTGGGTACCCGATGGA	182
QY	1344	LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly	1363
DB	183	AAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGAGACTACCTGTAAAGACGGT	242
QY	1364	AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn	1383
DB	243	GCCATAGCTTCAGATGCTGTGAGCTGCTTCCAGGATCACACTGTGATTTGAAC	302
QY	1384	IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn	1403
DB	303	ATCAATGATGTCAGTCTAATCCATGATGAAATCAGGCCACCTGTGTGATGATTAAT	362
QY	1404	SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer	1423
DB	363	TCATACAGTTGTAATGTCAGCCAGGATTTTCAGGCGMAAAGGTGTGMAACAGACTCT	422
QY	1424	ThrGlyPheAsnLeuAspGluValSerGlyIleTyrGlyTyrValMetLeuAspGly	1443
DB	423	ACAGGCTTTAACTCGATTTTGAAGTTTCTGCAATCTATGGATATGTCTATGCTAGATGC	482
QY	1444	MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspAspMet	1463
DB	483	ATGCTCCCATCTCTCATGCTTAACCTGTATCTTCTGATGATTAATCTCTGACACATG	542
QY	1464	AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu	1483
DB	543	AACTATGGAACACCAATCTCTATGATGATGATGATGATGATGATGATGATGATGATG	602
QY	1484	ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro	1503
DB	603	ACTGATTATACGGCTGGGTTCTTTATGTGAATGCGAGGAAAGATAAACAATGTTCCC	662
QY	1504	SerValAsnAspGlyArg	1509

Db 663 TCGTGTAATGATGGCAGA 680
RESULT 21
US-10-242-747-14
; Sequence 14, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13
; Prior Application Number: 09/764,881
; Prior Filing Date: 2001-01-17
; Prior Application Number: 60/179,065
; Prior Filing Date: 2000-01-31
; Prior Application Number: 60/180,628
; Prior Filing Date: 2000-02-04
; Prior Application Number: 60/214,886
; Prior Filing Date: 2000-06-28
; Prior Application Number: 60/217,487
; Prior Filing Date: 2000-07-11
; Prior Application Number: 60/225,758
; Prior Filing Date: 2000-08-14
; Prior Application Number: 60/220,963
; Prior Filing Date: 2000-07-26
; Prior Application Number: 60/217,496
; Prior Filing Date: 2000-07-11
; Prior Application Number: 60/225,447
; Prior Filing Date: 2000-08-14
; Prior Application Number: 60/218,290
; Prior Filing Date: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-747-14

Alignment Scores:
Pred. No.: 2,4e-119 Length: 680
Score: 1278.00 Matches: 224
Percent Similarity: 99.56% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 1
Query Match: 12.51% Indels: 0
DB: 16 Gaps: 0

US-09-977-053-6 (1-1842) x US-10-242-747-14 (1-680)

QY	1284	CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis	1303
DB	3	TGTGTTGATGTTGGCTATGTTGCACATGTTGAAGAGGATTTGTAGCCCTGCAT	62
QY	1304	CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu	1323
DB	63	TGTGAACAGAGAGTCAATGATGCAATGCTTAAATATGCACTGTGTGAA	122
QY	1324	AspGlnValGlyPheLeuCysValProGlyPheLeuGlyThrArgCysGly	1343
DB	123	GACCAAGTTGGGGATTCATGTGCAATGCCACCTGGATTTTGGGTACCCGATGGA	182
QY	1344	LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly	1363
DB	183	AAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGAGACTACCTGTAAAGACGGT	242
QY	1364	AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn	1383
DB	243	GCCATAGCTTCAGATGCTGTGAGCTGCTTCCAGGATCACACTGTGATTTGAAC	302
QY	1384	IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn	1403

Db 303 ATCAATGAATGTCAGTCTAATCCATGATAGAAATCAGGCCACCTGTTGGTGAATAAAT 362
Qy 1404 SerTyrSerCysLysCysGlnProGlnPheSerGlyGlnArgCysGluThrGluGlnSer 1423
Db 363 TCATACAGTTGTAATGTCAGCAGGATTTTCAGCGAAGAGTGTAACAGAACAGTCT 422
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrValMetLeuAspGly 1443
Db 423 ACAGGCTTTAACTGATTTTGAAGTTTCTGSCATCTATGATATGATGCTAGATGGC 482
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerAspMet 1463
Db 483 ATGCTCCACTCTCCATGCTCTAACTGATACCTTCTGGATGAATCTCTGACGACATG 542
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAsnGlySerAspAsnThrLeuLeu 1483
Db 543 AACTATGGAAACCAATCTCTATGCAAGTTGTAACGCGACGACATACCTTGTCTCTG 602
Qy 1484 ThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyValArgGluLysIleThrAsnCysPro 1503
Db 603 ACTGATTATTAACCGCTGGGTTCTTTATGATGATGGCAGGAAAGATACAACTGTCCC 662
Qy 1504 SerValAsnAspGlyArg 1509
Db 663 TCGGTGAATGATGCGAGA 680

RESULT 22

US-10-466-164-29
; Sequence 29, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAPFO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN, Jr., Edward H.;
; APPLICANT: PERALTA, Careyana H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT
; CURRENT APPLICATION NUMBER: US/10/466,164
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 3262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 LI:202943.4:2001JAN12
US-10-466-164-29

Alignment Scores:

Pred. No.: 4e-114 Length: 3262
Score: 1237.00 Matches: 240
Percent Similarity: 93.13% Conservative: 9
Best Local Similarity: 91.60% Mismatches: 9
Query Match: 12.10% Indels: 9
DB: 13 Gaps: 2

US-09-977-053-6 (1-1842) x US-10-466-164-29 (1-3262)

Qy 1590 LysSerLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpPro 1609
Db 289 GAGTCACTGCTACCTCTCTCCAGAGGAACTCAGTAAAGGAAACCTGTAGCAGTGGCT 348
Qy 1610 AspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysSerIlePheCysSer 1629
Db 349 GATTTCTTGTTCAGGAATTTGGGGGAAAGTGAAGATTCGATTCGAAGAGATATTTGTCT 408
Qy 1630 AspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLys 1649
Db 409 GGTTCGCCACGCTTAGAGGGGTGAGTGCCTCATCTGAGAACTGTCATCTGAAGATTTAAAG 468
Qy 1650 ProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProVal 1669
Db 469 CCAGGTTCCAAAGTCAATCTGTTCTGTGATCCAGGCTTCCAGCTGGTCCGGAACCTGTG 528
Qy 1670 GlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSer 1689
Db 529 CAGTACTGTCTGAATCAAGGACAGTGGACACAAACACTTCTCCTCACTGTGAACGCAATTAGC 588
Qy 1690 CysGlyValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGly 1709
Db 589 TGTGGGTGCCACTCTCTTGGAGATGGCTTCCATTCCAGCGGATGACTTCTATGCTGGC 648
Qy 1710 SerThrValThrTyrGlnCysAsnAsnGlyTyrTyrIleuLeuGlyAspSerArgMetPhe 1729
Db 649 AGCACAGTAACTACCACTAGTGCACAAATGGGCTACTTATTTGGGTGACTCAGGGATGTC 708
Qy 1730 CysThrAspAsnGlySer-TrpAsnGlyValSerProSerCysLeuAspValAspGluCys 1749
Db 709 GTACAGATATATGGAGCTAGGACGCGGTTTCCACCATCTCCGCTGATGTCGATGAGTG 768
Qy 1749 sAlaValGlySerAspCysSerGluHisAlaSerCysLeu-AsnValAspGlySerTyrI 1769
Db 769 TGCAGTTGGATCAGATTGTAGTGAGCATGCTTCTTGGCTGAAACGTAGATGGATCTTACA 828
Qy 1769 leCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysC 1789
Db 829 TATGTTTCATGTCTCCACCGTACACAGGAGATGGGAAAAAACTGTGCAGAACCTATAAAAT 888
Qy 1789 ys-LysAlaProGly---AsnProGluAsnGlyHisSerSerGlyGluIleTyrThr-Va 1807
Db 889 GCTAAGGCTCCAGCGCAGAAATCCGGAANAATGGCCACTCTCAGGTGAGATTTATACAGT 948
Qy 1807 lGlyAla---AlaValThrPheSerCysGlnGluGlyTyr-GlnLeuMetGlyValThrL 1826
Db 949 AGGTGCCGGAAGTCACATTTATTAGTGTACGAGGATACCCAGTTATGAGGATGTAACCA 1008
Qy 1826 ysIleThrCysLeuGlu-SerGlyGluTrpAsnHisLeuIlePro-TyrCys 1842
Db 1009 AAATCACATGTTGGAGTACTGGAGATGGAATCATCTAATACCAATATTGT 1060

RESULT 23

US-10-190-115-3
; Sequence 3, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgees, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.

; APPLICANT: Gusev, Vladimir Y.
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Lepley, Denise M.
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Mezick, Amanda J.
 ; APPLICANT: Padigar, Muralidhara
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Shen, Lei
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Spaderna, Steven K.
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Szekeres, Edward S. Jr.
 ; APPLICANT: Taupier, Raymond J. Jr.
 ; APPLICANT: Tchernev, Velizar T.
 ; APPLICANT: Zerhusen, Bryan D.
 ; APPLICANT: Voss, Edward Z.
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-050 CIP
 ; CURRENT APPLICATION NUMBER: US/10/190,115
 ; PRIOR FILING DATE: 2003-02-10
 ; PRIOR APPLICATION NUMBER: 60/303,168
 ; PRIOR FILING DATE: 2001-07-05
 ; PRIOR APPLICATION NUMBER: 60/368,996
 ; PRIOR FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: 60/386,816
 ; PRIOR FILING DATE: 2002-06-07
 ; PRIOR APPLICATION NUMBER: 60/215,854
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/215,856
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/215,902
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 60/216,585,
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/216,586
 ; PRIOR FILING DATE: 2001-07-07
 ; PRIOR APPLICATION NUMBER: 60/216,722
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/218,622
 ; PRIOR FILING DATE: 2000-07-17
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 136
 ; SOFTWARE: CuroseqList version 0.1
 ; SEQ ID NO 3
 ; LENGTH: 6728
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (13)..(4431)
 ; US-10-190-115-3

Alignment Scores:
 Pred. No.: 5,728-81 Length: 6728
 Score: 914.50 Matches: 355
 Percent Similarity: 34.59% Conservative: 174
 Best Local Similarity: 24.01% Mismatches: 529
 Query Match: 8.95% Indels: 550
 DB: 16 Gaps: 86

US-09-977-053-6 (1-1842) x US-10-190-115-3 (1-6728)

QY 281 CysAspGluGly-----LysAspCys 287
 DB 169 TGTCCAGAGGGCTTCTGGGGGAATATTGTCACATCGAGACCCCTGTGAGAGAACCGC 228
 QY 288 CysAspArgMetGlySerCysGlyThrHisThrGlyHisPheGluCysIleCys 307
 DB 229 TGCCAGAAATGGTGGGACTTGTGTGGCCAGGCCATGTCGGGGAAAGCCACGTCGCCGATGT 288
 QY 308 GluLysGlyTyTyTyGlyLysGlyLeuGlnTyGluCysThrAlaCysProSerGlyThr 327

DB 289 GCGTCAGGGTTACAGAGGAGGACTGCCAGTACTCA---ACATCTCATCATCCTTTGTG 345
 QY 328 TyrLysProGluGlySerProGlyGlyIleSerSerCysProCysProAspGluAsn 347
 DB 346 TCTCGACCC-----TGCGTG-----AATGGC 366
 QY 348 HisThrSerProGlySerThrSerProGluAspCysValCysArgGluGlyTyArg 367
 DB 367 GGCACATGCCATATGCTCAGCGGGANATACCTATGAGTCACCTGTCACTGGGGTTT--- 423
 QY 368 AlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysProProGluAsnGly 387
 DB 424 ---ACAGTAAGAGGTGCCAATGCGCGGATGCTGCTGTCTCATCTCTGCAATGGA 480
 QY 388 TyrPheIleGlnAsnThrCysAsn-----AsnHisPheAsnAlaCysGlyVal 404
 DB 481 ---AGTACCTGTACCACTGTGGCCACACAGTTCTCTGCAATGC----- 522
 QY 405 ArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsnGly 424
 DB 523 -----CTCAGGC----- 531
 QY 425 LeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGlnPro 444
 DB 532 ---TTCACAGGGCAGAAATGTCAGACTGATGTCATGAGTGTGACATTCACGACACTGC 588
 QY 445 LysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeuVal 464
 DB 589 CAGCATGGT-----GGCACCTGCCTC----- 609
 QY 465 AlaCysAspGluGlyTyArgLeuGluGlySerAspLysLeuThrCys---GlnGlyAsn 483
 DB 610 -----AACCTGCTGGTTCTACCACTGCGCAGTGCCTCTCAGGGC--- 648
 QY 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
 DB 649 -----TTCACAGGCCAGTACTGTGACAGCTGTATGTG 681
 QY 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
 DB 682 CCC-----TGTCACCTCCACCTTGTGTCAATGAGGC 714
 QY 524 IleCys-----TyrValSerCysArgGlnGlyPheIleLeu 535
 DB 715 ACCTGTGGCAGACTGGTGACTTCACTTTTGTAGTGCACCTGCTCCAGGTTTGAAGG 774
 QY 536 SerGlyValLysGluMetLeu-----ArgCysThrThrSerGlyLys 549
 DB 775 AGCACCTGTGAGAGGAATATTGATGACTGCCCTAACCCACAGGTGTGAGATGAGGGTT 834
 QY 550 TrpAsnValGlyValGlnAlaValCysLysAspValGluAlaProGlnIleAsn--- 568
 DB 835 TGTGTGGATGGGGTCAACACTTACACTGCCCTGT-----CCCCACAAATGACGAGGA 888
 QY 569 -----CysProLysAspIleGluAlaLysThrLeuGlnGlnAspSerAlaAsnVal 586
 DB 889 CAGTTCTGACAGAGAGTGGATGAATGCCCTGCTGCAG----- 927
 QY 587 ThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerValHisValHis 606
 DB 928 -----CCCAATGCCCTGTCAAAATGGGGC----- 951
 QY 607 ProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleValTyrThrAla 626
 DB 952 -----ACCTGT 957
 QY 627 ThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIleAspAlaGlu 646
 DB 958 GCCACCGCAATGAGGCTAT---GCCTGTGTATGT----- 990
 QY 647 ProProValIleAspTrpCysArgSerProProValGlnValSerGluLysValHis 666

1331 slyeCysProProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSe 1351
1351 rGlnProCysLysAsnGlyAlaThrCysLysAsnGlyAlaAsnSerPheArgCysLeuCy 1371
2703 TGAATGTCACAGGCTTCAGTGATGAGTCTGTCAGGAGGACATTCATGATGATGCTTTCG 2762
1371 rGlnProCysLysAsnGlyAlaThrCysLysAsnGlyAlaAsnSerPheArgCysLeuCy 1371
2763 CAATCTTCGAGGAGGAGGCTTCATGATGAGTGAATGATGATGATGATGATGATGATG 2822
1371 sAlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnPr 1391
2823 CTTTCGGGTTTCAGTGGGATAAGTCCAGACACACATGATGATGATGATGATGATGATG 2882
1391 oCysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnPr 1411
2883 CTGTAAAGATGAGGAGGAGGCTTCATGATGATGATGATGATGATGATGATGATGATG 2942
1411 oGlyPheSerGlyGlnArgCysGluThrGlu-----GlnSerThrGlyPh 1426
2943 AGGATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3002
1426 eAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAsp-GlyMetLeuP 1446
3003 CAAT-----GGTGGCACATGCTGTTGATGGATTAACATC 3035
1446 roSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspAspMetAsnTyrG 1466
3036 CTTCCT-----CTTGCTTGCCCTGTGGGTT----- 3061
1466 lyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspT 1486
3062 -----TCACTGGATTCCTTCCTCCCTCCATGATCA 3092
1486 yrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysProSerVala 1506
3093 TGAATGACGCTCTCATCCATGCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3130
1506 sAspGlyArgTyrHisIleAlaIleThrTyr-----ThrSerAlaAsnGlyIleT 1524
3131 -----GTGTTGATGGCTGGGTACCTACCTCCCTGCGAGTGGCCCT 3170
1524 rPLysValTyrIleAspGlyLysLeuSerAspGlyGlyValaGlyLeuSerValGlyLeu 1543
3171 GGGCTACAGT-----GGAAAACTGTGAGACCTGGTGAATCTCTGCA-GTGGGTCT- 3222
1544 ProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysLysGlyGlu---- 1562
3223 CCATGTAAAAACAAGGTACTTGTGT-----CAGAAAAAGCAGAGTCC 3267
1563 -----GlyPheSerProAlaGluSerPheValGlySerIleSerGln 1576
3268 CAGTGCCTATGTCATGATGGCTGGTGGCTATTTGTGACGTGCCCAATGTC----- 3321
1577 LeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCys 1596
3322 -----TCTGT 3327
1597 ProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleVal 1616
3328 GACATAGCAGCTCCAGGAGAGGTGCTGTT-----GAACACTTGTGCCAGCACTCA 3381
1617 GlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGly 1636
3382 GGT---GTCTGATCAATCTGCAACAGCATTTACTGTGAGTGCCCTC-----CTGGGCTAT 3435
1637 SerValProHisLeu-----ArgThrAlaSerGluAspLeuLysProGly 1651
3436 ACTGGAGGACTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3495
1652 SerLysValAsnLeuPhe-----CysAspProGlyPheGlnLeu 1664
3496 GCAACATGAGTCACTTCATTGGTGTGATACAGATGGAGTGTGTGCCAGGCTATCAGGCT 3555
1665 ValGlyAsnProVal-----GlnTyrCysLeuAsnGlnGlyGln 1677

3556 GTCAACTGTGAGTATGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 3615
1678 TrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGlu 1697
3616 TGTATTGACCTTGTGAACCAT-----TTCAAGTGTCTTGTGCCACCA----- 3657
1698 AsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsn 1717
3657 ----- 3657
1718 AsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsn 1737
3658 -----GGCACTGGGGCTTACTCTGTGAAGAGAAC----- 3687
1738 GlyValSerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGlu 1757
3688 -----ATTGATGACTGTGCGCGGCTGCCCATTCCTTAAT 3723
1758 HisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThr 1777
3724 GGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3783
1778 GlyAspGlyLysAsnCysAlaGluProIleLys-----CysLysAla 1791
3784 GGGGAG-----CGTTGTGAGGAGACATCAAGAGTGTCTTCCAAACCTGAGCTCT 3837
1792 ProGlyAsn 1794
3838 GAGGGCAGC 3846

RESULT 24
US-10-369-072-3
; Sequence 3, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Szytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosche, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369, 072
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174, 372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898, 994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215, 854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215, 856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215, 902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216, 585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216, 586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216, 722
; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 6728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-369-072-3

Alignment Scores:
Pred. No.: 5,72e-81 Length: 6728
Score: 914.50 Matches: 395
Percent Similarity: 34.59% Conservative: 174
Best Local Similarity: 24.01% Mismatches: 529
Query Match: 8.95% Indels: 550
DB: 16 Gaps: 86

US-09-977-053-6 (1-1842) x US-10-369-072-3 (1-6728)

```
QY 281 CysAspGluGly-----LysAspCys 287
DB 169 TGTCCAGAGGCTTCTGGGGGAATATTGCAACATCGAGACCCCTGTGAGAGAACCGC 228
QY 288 CysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPheGluCysIleCys 307
DB 229 TGCAGAAATGTGGGAGTCTGTGTGGCCAGGCCATGCTGGGGGAAGCAAGTGTCCCATGT 288
QY 308 GluLysGlyTyrTyrGlyLysGlyLeuGluTyrGluCysThrAlaCysProSerGlyThr 327
DB 289 GCCTCAGGGTTTACAGAGAGAGACTGCCAGTACTCA---ACAICTCATCCATGCTTTGTG 345
QY 328 TyrLysProGluGlySerProGlyGlyIleSerSerCysIleProCysProAspGluAsn 347
DB 346 TCTCGACCC-----TGCCTG-----AATGGC 366
QY 348 HisThrSerProGlySerThrSerProGluAspCysValCysArgGluGlyTyrArg 367
DB 367 GGCACATGCCATATGCTCAGCGGATACCTATGAGTGCACCTGTCAAGTCGGGTTT--- 423
QY 368 AlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysProProGluAsnGly 387
DB 424 ---ACAGGTAAAGAGTGCCTCAATGACGAGTGCCTGCTCATCCCTGTGCATATGGA 480
QY 388 TyrPheileGlnAsnThrCysAsn-----AsnHisPheAsnAlaCysGlyVal 404
DB 481 -----AGTACCTGTACCATCTGTGGCCACAGTCTCTCTGCAATGC----- 522
QY 405 ArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCysLeuProAsnGly 424
DB 523 -----CTCAGAGC----- 531
QY 425 LeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGlnPro 444
DB 532 ---TTCCAGAGGGCAGAAATGTGAGACTGATGTCAATGAGTGCACATTCCTCAGGACACTGC 588
QY 445 LysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrCysLeuVal 464
DB 589 CAGCATGGT-----GGCAGCTGCTCTC----- 609
QY 465 AlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys---GlnGlyAsn 483
DB 610 -----AACCTGCTGTGTTCTTACAGTGCAGGCTGCTGCTCAGGCGC--- 648
QY 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
DB 649 -----TTCCAGGCGCAGTACTGTGACAGCTGTATGTG 681
QY 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
```

```
DB 682 CCC-----TGTGCACCCCTCACCTTGTGTCAATGGAGGC 714
QY 524 IleCys-----TyrValSerCysArgGlnGlyPheIleLeu 535
DB 715 ACCTGTGCGCAGACTGTGTGACTTTCACCTTTGAGTGCACTCCCTTCAGGTTTGAAGG 774
QY 536 SerGlyValLysGluMetLeu-----ArgCysThrThrSerGlyLys 549
DB 775 AGCACCTGTGAGAGGAATATTGATGCTGCCCTAACACACAGGTGTGACAAATGGAGGGTT 834
QY 550 TrpAsnValGlyValGlnAlaValCysLysAspValGluAlaProGlnIleAsn--- 568
DB 835 TGTGTGATGGGTCAACACTTCACTGCGCTGT-----CCCCACAAATGGACAGCA 888
QY 569 -----CysProLysAspIleGluAlaLysThrLeuGluGlnGlnAspSerAlaAsnVal 586
DB 889 CAGTTTGCACAGAGGATGGATGATGCTGCTGCAG----- 927
QY 587 ThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerValHisValHis 606
DB 928 -----CCCAATGCTGTCAAAATGGGGC----- 951
QY 607 ProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleValTyrThrAla 626
DB 952 -----ACCTGT 957
QY 627 ThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIleAspAlaGlu 646
DB 958 GCCAACCCCAATGAGGCTAT---GGCTGTGTATGT----- 990
QY 647 ProProValIleAspTrpCysArgSerProProValGlnValSerGluLysValHis 666
DB 991 -----GTCAACGGCTGAGTGGAGAT-----GACTGCAGTGCAGACATT--- 1029
QY 667 AlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeuValIleThr 686
DB 1030 -----GATGATGTGCTTGCCTTCTGCTACTCCAGGCTCCACCTGCATCGAC 1077
QY 687 ArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGlnTyrThrAla 706
DB 1078 CGTGTGCGCTCTCTCTCTGATGTGCCAGAGGGAGGAGGAGGCTCTCTGTGTCACTG 1137
QY 707 ThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySerPro 726
DB 1138 GATGAT-----GCATGCATCATCAATCTTCCACAGGGGCACTG 1179
QY 727 CysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAspAsnThrGly 746
DB 1180 TGTGACACC-----AACCCCTAAATGGGCAATATATTGC----- 1215
QY 747 ValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySerThrAspLys 766
DB 1216 -----ACCTGCCACAGGCTTACAAAGGGCTGCTGTCACAGAGATGTG 1260
QY 767 TyrTyrCysAlaTyr-----Glu 772
DB 1261 GATCAATGTGCCATGGCCATAGCAATCTTGTGAGCATGTCAGGAAATGTGTGACACG 1320
QY 773 AspGlyValTrpLysProThrTyrThrThrGluTrpProAspCysAlaLysLysArgPhe 792
DB 1321 GATGGCGCTTCCAC-----TGTGAGTGTCTCAAGGGT----- 1353
QY 793 AlaAsnHisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThr 812
DB 1354 -----TATGCAGGACCTCGTTGT---GAGATG 1377
QY 813 AspLeuMetLysLysPheSerGluAlaPheGlu-----ThrThrGluLysMet 829
DB 1378 GACATCATGATGCCCATTCAGACCCCTGCCAGATGATGCTACTGCTCGATAGATT 1437
QY 830 -----ValProSerPheCysSerAspAlaGluAspIleAspCys 842
DB 1438 GGAGGCTTCATGTCTGTGATGCCAGGTTTC-----AAAGGTGTGATTTGT 1485
```


Db 277 GCCTCAGGTTTACAGGAGAGACTGCCAGTACTCG---ACATCTCATCCATCGCTTTGTG 333
Qy 328 TyrLysProGluGlySerProGlyGlyIleSerCysIleProCysProAspGluAen 347
Db 334 TCTCAGCT-----TGCCTG-----AATGGC 354
Qy 348 HisThrSerProGlySerThrSerProGluAspCysValCysArgGluGlyTyrArg 367
Db 355 GGCACATGCATATGCTCAGCCGGATACCTATGAGTGCACCTGTCAAGTCGGT--- 411
Qy 368 AlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysProGluAenGly 387
Db 412 ---ACAGGTAGAGTGC---AATGCGAGTGCCTGCTCTCATCCCTGTGCCAATGGA 468
Qy 388 TyrPheIleGlnAsnThrCysAen-----AsnHisPheAsnAlaAlaCysGlyVal 404
Db 469 -----AGTACCTGTACCACTGTGGCCACACAGTCTCTCTGCAATGC----- 510
Qy 405 ArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAenGly 424
Db 511 -----CTCAGGC----- 519
Qy 425 LeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGlnPro 444
Db 520 ---TTCACAGGCGAGAAATGTCAGACTGATGCAATGAGTGTGACATTCACGACACTGC 576
Qy 445 LysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrCysLeuVal 464
Db 577 CAGCATGGT-----GGCACCTGCCTC----- 597
Qy 465 AlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys---GlnGlyAen 483
Db 598 -----AACCTGCTGCTTCTACCACTGCAGTGCAGTGCCTCAGGCG--- 636
Qy 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
Db 637 -----TTCACAGGCGAGTACTGTGACAGTGCAGCTGTATGTG 669
Qy 504 ProLysAspValIleIleSerProHisAenCysGlyLysGlnProAlaLysPheGlyThr 523
Db 670 CCC-----TGTGCACCCCTCACCTTTGTGTCAATGGAGCG 702
Qy 524 IleCysTyrValSerCysArgGln-----Gly 532
Db 703 -----ACCTGTGCGCAGACTGCTGACTTCACTTTTGTGAGTGCATTTACCAAGT 750
Qy 533 PheIleLeuSerGlyValLysGluMetLeu-----ArgCysThrThr 546
Db 751 TTTGAGGGAGACCTGTGAGAGGAATATTGATGACTGCCCTTAACACAGGTGTGAGAT 810
Qy 547 SerGlyLysTrpAenValGlyValGlnAlaAlaValCysLysAspValGluAlaProGln 566
Db 811 GGAGGGTTTGTGTGATGGGTCAACACTTCAACACTGCCGTG-----CCCCACAA 864
Qy 567 IleAen-----CysProLysAspIleGluAlaLysThrIleuGluGlnAspSer 583
Db 865 TGCACAGGACACTTCTGCACAGAGATGTGGATGAATGCTCTG-----CAGCCCAAT 918
Qy 584 AlaAenValThrTrpGlnIleProThrAlaLysAspAenSerGlyGluLysValSerVal 603
Db 919 GCCTGTCAAACTGG-----GGCACCTGTGCCAACCGCAATGGAGGCTATGGCTGTGTA 972
Qy 604 HisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleVal 623
Db 973 TGTGTCAACGGC----- 984
Qy 624 TyrThrAlaThrAspLeuSerGlyAenGlnAlaSerCysIlePheHisIleLysValIle 643
Db 985 TGGAGTGGAGATGACTGCACTGAGAACATTTGATGATGATGCTTTCG---CGCGCTGTACTC 1042
Qy 644 AspAlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGlu 663

Db 1043 CAGGCTCCACCTGCATCGACCGTGTGG-----CCTCCTTCTCTTGCATGTGCCAG 1093
Qy 664 LysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAenSerGlyAlaGluLeu 683
Db 1094 AGGGGAAGCGAGGTCTCC-----TGT 1114
Qy 684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
Db 1115 GTCATCTGATGATCATGTCATGCAATCTTGGCCAAAGGGGGCAC----- 1162
Qy 704 TyrThrAlaThrAspProSerGlyAsnAenArgThrCysAspIleHisIleValIleLys 723
Db 1162 ----- 1162
Qy 724 GlySerProCysGluIleProPheThrProValAenGlyAspPheIleCysThrProAsp 743
Db 1163 -----TGTGTGACACCA---ACCCCTTA---AATGGCAATATATATTTGC----- 1200
Qy 744 AsnThrGlyValAenCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
Db 1201 -----ACCTGCCACAGGCTACAAAGGGGCTGACTGCACA 1236
Qy 764 ThrAspLysTyrTyrCysAlaTyr----- 771
Db 1237 GAAGATGTGATGATGTGCCATGCGCAATAGCAATCTTGTGAGCATGCGAGAAATGT 1296
Qy 772 -----GluAspGlyValTrpLysProThrTyrThrGluTrpProAspCysAlaLys 789
Db 1297 GTGAACACGGATGGCGCTTCCAC-----TGTGAGTCTCTGAAG 1335
Qy 790 LysArgPheAlaAenHisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCys 809
Db 1336 GGT-----TATGCAGGACCTCGTTGT 1356
Qy 810 AspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu-----ThrThrLeu 826
Db 1357 ---GAGATGGACATCAATGAGTGCCTTCCAGACCCCTGCAGATGATGTACTGTCTG 1413
Qy 827 GlyLysMet-----ValProSerPheCysSerAspAlaGluAsp 839
Db 1414 GATAAGATTGGAGGCTTCATGTCTGTGTCATGCCAGGTTTC-----AAAGGT 1461
Qy 840 IleAspCysArgLeuGluAen-----LeuThrLysLysTyr 852
Db 1462 GTGCATTGTGAATTAATAATAATGATGTCAGACCACTTGTGTGACAAATGGCGAG 1521
Qy 853 CysLeuGluTyrAsnTyrAspTyrGluAenGlyPhe-----AlaIleGlyProGlyGly 870
Db 1522 TGTGTG-----GATAAAGTCAATCGTTTCCAGTGTGTGCTCTCTCTGTTTC 1569
Qy 871 TrpGlyAlaAlaAenArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGlu 890
Db 1570 ACTGGGCCAGTTTGCAGATTGAT-----ATTGATGACTGT----- 1605
Qy 891 ThrAlaThrSerIleGlyAen-AlaLysSerSerArgIleLysArgSerAlaProLeuSe 910
Db 1606 TCCAGTACTCCCTGTCTGAATGGGCAGAGTATGATGATCACCAGATGGCT----- 1657
Qy 910 rAspTyrLysIleLysLeuIlePheAenIleThrAlaSerValProLeuProAspGluAr 930
Db 1658 -----ATGAATGCCAGTGTGCCA----- 1675
Qy 930 GAsnAspThrLeuGluTrpGluAenGlnArgLeuLeu-----GlnThrLe 946
Db 1676 -----CAGTTTTCATGGTGTGTGTGTGAGGAGACATTT 1710
Qy 946 uGluThrIleThrAsnLysLeuLysArgThrLeuAenLysAspProMetTyrSerPheGl 966
Db 1711 GACAACTGTGACCCC-----GATCCTTGGCCACCATGTCTCA 1745
Qy 966 nLeuAlaSerGluIleLeuIleAlaAspSerAenSerLeuGluThrLysLysAlaSerPr 986
Db 1746 GTGTCCAGGATGGTATT-----GATTCCTACACCTGC----- 1776


```
Db 3532 TATCAGGGTGTCAACTGTGAGTATGAGTGCAGATCACCCCTGCCAGAT 3591
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1676 GlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProPro 1695
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3592 GGAGGCACCTGTATTGACCTTTGTGAACCATTTCAAGTGTCTTGC-----CCACCA 3642
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1696 LeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGln 1715
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3643 GGCACCTCGGGGTATGAATCA-----TCCTTATCCATTTCAT 3681
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1716 CysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySer 1735
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3682 TGCCGGGTCCCATGTGCTTAATGGTGTGAG-----TGCATGGTAGG----- 3726
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1736 TrpAsnGlyValSerProSerCysLeu-----Asp 1745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3727 ATTGGAGGCTACAGTTGTGCTGCTTGTGCTTGTGGGGAGCGTTGTGAGGGAGAC 3786
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1746 ValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer-----CysLeuAsn 1763
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3787 ATCAACGAGTGC---CTCTCCAAACCCCTGCAGCTCTGAGGGCAGCCTGGACTGTATACAG 3843
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGly----- 1778
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3844 CTCACCAATGACTACTGTGTGTGTTCCTGCTAGTGCCTTTACTGTGGGCACTGTGAAACC 3903
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1779 -----AspGlyLysAsnCysAlaGlu-ProIl 1787
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3904 TTCGTGATGTGTGCCAGATGCCCTGCCTGTAATGGAGGAGCTTGTGCTGTGGCCAGT 3963
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1787 eLysCys 1789
      : : : : :
Db 3964 AACATGC 3970
```

Search completed: May 12, 2004, 00:23:10
Job time : 2648.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.

Run on: May 10, 2004, 15:48:01 ; Search time 13847.8 seconds
(without alignment)
3972.182 Million cell updates/sec

Title: US-09-977-053-6
Perfect score: 10219
Sequence: 1 MWFLAPCCWGLALVSGMAT.....GVTKITCLSEGMNHLIPYC 1842

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US9977053/runat_06052004_075943_18200/app_query.fasta_1.5710
-DB=EST -OPMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=25 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9977053@cgn_1_15412 @runat_06052004_075943_18200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database :

RST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpi: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_rod: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gse1: *

29: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5409.5	52.9	3288	29	AY405591	AY405591 Homo sapi
2	5060.5	49.5	3009	29	AY405592	AY405592 Pan trogl
3	4803.5	47.0	3288	29	AY405593	AY405593 Mus muscu
4	4781	46.8	3171	11	BC030816	BC030816 Homo sapi
5	3955	38.7	2736	11	AK052699	AK052699 Mus muscu
6	3011	29.5	3127	14	CB548140	CB548140 AMGNNUC:C
7	2023.5	19.8	3242	11	AK035333	AK035333 Mus muscu
8	1793	17.5	3307	11	AK052463	AK052463 Mus muscu
9	1690	16.5	906	13	BX390270	BX390270 BX390270
10	1470.5	14.4	1201	13	EX356254	EX356254 BX356254
11	1407.5	13.8	1201	13	EX356253	EX356253 BX356253
12	1385	13.6	748	13	EX093982	EX093982 BX093982
13	1359	13.3	738	9	ALU706149	ALU706149 DKFZP686E
14	1304	12.8	801	9	ALU137280	ALU137280 AU137280
15	1295	12.7	756	12	BG620159	BG620159 602618422
16	1188	11.6	829	12	BG434540	BG434540 602506592
17	1176.5	11.5	701	13	B0929100	B0929100 AGENCOURT
18	1145	11.2	719	12	BG436029	BG436029 602508809
19	1125.5	11.0	1887	11	AK045217	AK045217 Mus muscu
20	1050	10.3	2408	11	AK087058	AK087058 Mus muscu
21	1039.5	10.2	2019	11	BC047729	BC047729 Homo sapi
22	1027.5	10.1	1035	12	BG621817	BG621817 602619678
23	1022	10.0	642	10	BBS00305	BBS00305 BBS00305
24	990	9.7	673	10	BBS42219	BBS42219 BBS42219
25	982.5	9.6	998	13	BQ17243	BQ17243 AGENCOURT
26	982	9.6	763	13	BQ392790	BQ392790 603805162
27	958	9.4	652	13	BQ367139	BQ367139 MR3-GN018
28	922	9.0	503	10	BF367968	BF367968 RC3-GN004
29	898.5	8.8	482	10	BF998780	BF998780 QV0-GN023
30	875.5	8.6	9330	11	AK090118	AK090118 Mus muscu
31	856	8.4	480	10	BG003364	BG003364 RC3-GN004
32	843	8.2	579	10	BF367305	BF367305 MR0-GN002
33	759	7.4	477	10	BG012274	BG012274 RC2-GN028
34	757.5	7.4	5925	29	AY418022	AY418022 Mus muscu
35	745	7.3	415	10	BG005173	BG005173 CM2-GN022
36	739.5	7.2	4131	29	AY418953	AY418953 Homo sapi
37	735.5	7.2	444	10	BG012481	BG012481 RC3-GN027
38	734	7.2	412	10	BF369449	BF369449 RC0-GN009
39	730	7.1	734	13	BQ327243	BQ327243 603493537
40	728	7.1	433	10	BF999622	BF999622 RC2-GN013
41	721	7.1	4128	29	AY418955	AY418955 Mus muscu
42	718	7.0	452	13	BF693745	BF693745 BY693745
43	718	7.0	3721	11	BC032414	BC032414 Homo sapi
44	713	7.0	398	10	BG009206	BG009206 IL5-GN024
45	711.5	7.0	806	12	BG782985	BG782985 SEAMUC002
46	708	6.9	450	10	BG012275	BG012275 RC2-GN028
47	704	6.9	770	12	BG619827	BG619827 602618001
48	690	6.8	358	10	BF771579	BF771579 IL5-IT002
49	689	6.7	6838	29	AY418021	AY418021 Pan trogl
50	688	6.7	798	12	BG782977	BG782977 SEAMUC002
51	684	6.7	392	10	BF994869	BF994869 QV0-GN021
52	683	6.7	6848	29	AY418020	AY418020 Homo sapi
53	675.5	6.6	369	10	BH181604	BH181604 CM3-HT063
54	670	6.6	770	12	BG780526	BG780526 SEAMUC000
55	665	6.5	412	14	H72144	H72144 Y+99C05.r1
56	658	6.4	382	10	BF368046	BF368046 RCS-GN004
57	657.5	6.4	837	12	BG786173	BG786173 SEAMUC006
58	654	6.4	488	14	R92857	R92857 YQ13e07.r1
59	652.5	6.4	410	14	H72132	H72132 Y+99a05.r1
60	646	6.3	477	13	BQ366017	BQ366017 RC0-GN009
61	645.5	6.3	4131	29	AY418954	AY418954 Pan trogl
62	635.5	6.2	423	13	BQ366619	BQ366619 MR0-GN017

63	631.5	6.2	1815	29	AV398787	AV398787 Homo sapi
64	630	6.2	792	14	CD351453	CD351453 UI-M-G10-
65	629	6.2	442	13	BY360080	BY360080 BY360080
66	626.5	6.1	723	14	CF540496	CF540496 UI-M-GV0-
67	623	6.1	334	10	BF368870	BF368870 UC6-GN007
68	622.5	6.1	1815	29	AV398789	AV398789 Mus muscu
69	618	6.0	2098	11	BC032053	BC032053 Homo sapi
70	614	6.0	755	14	CD351379	CD351379 UI-M-G10-
71	611	6.0	848	13	EX728191	EX728191 BX728191
72	610	6.0	367	14	R67816	R67816 Y124f10.r1
73	607.5	5.9	671	14	CD351340	CD351340 UI-M-G10-
74	605	5.9	974	14	CF592712	CF592712 AGENCOURT
75	600	5.9	422	14	R33004	R33004 Yh7h12.r1
76	600	5.9	805	14	CF537892	CF537892 UI-M-G10-
77	599.5	5.9	381	14	R30986	R30986 Yh60d01.r1
78	598	5.9	733	14	CD803516	CD803516 UI-M-GV0-
79	597	5.8	336	14	HL3496	HL3496 Y115f07.r1
80	595	5.8	362	10	BF997406	BF997406 PML-GN018
81	591	5.8	3392	11	AK044597	AK044597 Mus muscu
82	589.5	5.8	779	12	BY523794	BY523794 603052095
83	588.5	5.8	342	10	BG008596	BG008596 PM4-GN030
84	588	5.8	658	9	AL778052	AL778052 AL778052
85	581	5.7	3394	11	AK004739	AK004739 Mus muscu
86	579	5.7	657	13	BU705790	BU705790 UI-M-F00-
87	578.5	5.7	1013	13	BU519205	BU519205 AGENCOURT
88	577.5	5.7	742	14	CF744242	CF744242 UI-M-GV0-
89	575.5	5.6	427	10	BG009198	BG009198 RCL-GN019
90	573	5.6	772	14	CB524782	CB524782 UI-M-FY0-
91	573	5.6	834	14	CB523447	CB523447 UI-M-GH0-
92	572	5.6	589	10	BF422341	BF422341 Lr CA12.0
93	571	5.6	310	10	BE181601	BE181601 CN3-HT063
94	571	5.6	1201	13	EX439805	EX439805 BX439805
95	570.5	5.6	1201	13	EX420707	EX420707 BX420707
96	570	5.6	311	12	BI051224	BI051224 CM4-GN036
97	570	5.6	537	13	EX301808	EX301808 BX301808
98	570	5.6	537	13	EX303884	EX303884 BX303884
99	570	5.6	537	13	EX303885	EX303885 BX303885
100	570	5.6	930	13	EX327882	EX327882 BX327882

RESULT 1	AY405591	3288 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	Homo sapiens HCM2275 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
DEFINITION	AY405591.1 GI:39761565				
ACCESSION	AY405591				
VERSION	GSS.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
ORGANISM	1 (bases 1 to 3288)				
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
TITLE	Science 302 (5652), 1960-1963 (2003)				
JOURNAL	14671302				
PUBMED	2 (bases 1 to 3288)				
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	Direct Submission				
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
JOURNAL	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
COMMENT					

FEATURES	Location/Qualifiers
source	1..3288
gene	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	-1..->3288
	/locus_tag="HCM2275"
ORIGIN	
Alignment Scores:	
Pred. No.:	0
Score:	5409.50
Percent Similarity:	90.42%
Best Local Similarity:	90.33%
Query Match:	52.94%
DB:	29
	Length: 3288
	Matches: 1000
	Conservative: 1
	Mismatches: 87
	Indels: 19
	Gaps: 1
US-09-977-053-6 (1-1842) x AY405591 (1-3288)	
QY	1 MetTTPProArgLeuAlaPheCysCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr 20
DB	1 ATGTGGCTCGCTGGCTGGCTTTTGTCTGGGGTCTGGCGCTCGTTTGGGGCTGGCGACC 60
QY	21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
DB	61 TTTCAGCAGATGTCCTGGCGCAATTTTCAGCTTCGGCTNNNNNNNNNNNNNNNNNN 120
QY	41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaGlySerArg 60
DB	121 NNN 180
QY	61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
DB	181 NNN 240
QY	81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg 100
DB	241 CGCTGGAGCTTGTCTCTCTCTGGTGGATGATTCGTCCAGCGTGGCGGAGTCACTTCGC 300
QY	101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
DB	301 AGCGAGCTCATGTTCTGTCGCAAGCTGCTGTCGACTTCTCCCGTGGTGGCCACG 360
QY	121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
DB	361 CGCGTGGCCATCGTGACCTTCTCGTCCAAGAACTACGTGGTGGTGGCGCGTCAATCATC 420
QY	141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGlnIleProAlaIle 160
DB	421 TCACCCCGCGCGCGCGCGCAAGTGGCGCTGCTCTCCAGAGATCCCTGCCATC 480
QY	161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeu 180
DB	481 TCTACCGAGGTGGCGGCACTTACCAAGGGCGCTTCCAGCAAGCGCGCAAAATCTT 540
QY	181 LeuHisAlaArgGluAsnSerThrLysValValPheIleThrAspGlyTyrSerAsn 200
DB	541 CTTTCATGCTAGAGAAAACCAACAAAGTTGTATTTCTCATCTCATGTGATATTCAT 600
QY	201 GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
DB	601 GGGGGAGACCTTAGACCAATTCAGCGTCACTCGAGATTCAGAGATTCCTCACT 660
QY	221 PheGlyIleTTPGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
DB	661 TTTTGCATATGGCAAGGAACATTCAGAGCTCAATGACATGGCTTCCACCCCAAGAG 720
QY	241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeu 260
DB	721 GAGCACTGTGTACTGCTACACAGTTTGAAGAAATTTAGAGCTTTAGCTCCCGGGCAT 780
QY	261 HisGluAspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyrLeu 280

Db 781 CATGAAGATCTACCTTCGGGAGCTTTTATTCAAGATCATATGCTCCACTCATATCTT 840
 Qy 281 CysAspGluGlyAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
 Db 841 TGTGATGAAGGCAAGACTGCTGTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACACA 900
 Qy 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
 Db 901 GGCCATTTTGAGTGCATCTGTGAAGAGGGTATTACGGGAAGGTCTTCAGTATGAATGC 960
 Qy 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
 Db 961 ACAGCTTGCCCATCGGGACATACAAACCTGGAAGGCTCACACAGAGCAATCAGCAGTTC 1020
 Qy 341 IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys 360
 Db 1021 ATTCCATGTCTGATGANAATCACCTCTCCACCTGGGAAGGCATCCCTGGAAGACTGT 1080
 Qy 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
 Db 1081 GTCTGCGAGAGAGGATACAGGGGCACTTGGCCAGACTGTGTCACTTGTCCACTGCCCTGGC 1140
 Qy 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
 Db 1141 CTGAAGCCTCCCGAANAATGTTACTTTATCCAAAACACTTGCACCAACCACTTCAATGCA 1200
 Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
 Db 1201 GCCTGTGGGGTCCGATGTCCCTCGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1260
 Qy 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
 Db 1261 CTACCCCATGGTTTGTGTCGGTTTCAGAGAGTACTTCGAGAGTAGAAGCATGTCTCAT 1320
 Qy 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
 Db 1321 CTCGGCAGCCGAAACATGCCCACATCAGCTGTCTACAGGGCAATGTTATATAAGACA 1380
 Qy 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
 Db 1381 ACATGTTTGGTTGCTGTGATGAAGGGTACAGACTAGAAGGCCAGTGAATAGCTTACTTGT 1440
 Qy 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
 Db 1441 CAAGGAACACCGGTGGTGGAGTGGCCAGAACCCCGGTGTGTGGAGGCCACTGTTCACC 1500
 Qy 501 PheGlnMetProLysAspValIleLeSerProHisAsnCysGlyLysGlnProAlaLys 520
 Db 1501 TTTTCAGATGCCCAAGATGTTCATCATATATCCCCCACTGTGGCAAGCAGCCAGCCAAA 1560
 Qy 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
 Db 1561 TTTGGGACGATCTGCTATGTATGTAGTGGCCGCAAGGGTTCATTTTATCTGGAGTCAAGAA 1620
 Qy 541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLys 560
 Db 1621 ATGCTGAGATGTACCACTTCTGGGAATGAATGTGAGTTCAAGCAGCTGTGTAAA 1680
 Qy 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
 Db 1681 GAGTGGAGGCTCTCTCAATCAACTGTCTTAAGGACATAGAGGCTAAGACTCTGGAACAG 1740
 Qy 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
 Db 1741 CRAGATTTCTGCCAATGTATTCTGGCAGATTCCAAACAGCTAAAGCAACTCTGGTGAAG 1800
 Qy 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
 Db 1801 GTGTGAGTCTCAGCTTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGT 1860
 Qy 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
 Db 1861 GCTATCGTATACAGGCNACTGACCTATCCGGCAACCCAGGCCAGCTGCGTGTTCCTATC 1920

Qy 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProProValGln 660
 Db 1921 AAGGTTATTGATCGAGAACCACTGTCTATAGACTGTGTGCAGATCTCCACTCCGGTCCAG 1980
 Qy 661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
 Db 1981 GTCTCGCAGAAGGTACATCGCGCAAGCTGGCATGAGCTCAGTCTCTCAGACAACTCAGG 2040
 Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
 Db 2041 GCTGAATTTGCTCATTACCAGAGTCATACCAAGGAGACCTTTTCCCTCAAGGGGAGACT 2100
 Qy 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
 Db 2101 ATAGTACGATATACGCCACTGACCCCTCAGGCCAATACAGGCATGTGATATCCATAT 2160
 Qy 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
 Db 2161 GTCATAAAA----- 2169
 Qy 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
 Db 2170 -----GATAATATCTGGAGTCAACTGTACATTAACCTTGTGGAGGCTATGATTTTCA 2223
 Qy 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
 Db 2224 GAAGGCTTACTGACAGATTAATTGTGCTTATGAGATGGCGTCTGGAAACCAACATAT 2283
 Qy 781 ThrThrGlnTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
 Db 2284 ACCACTGAATGGCCAGACTGTGCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2343
 Qy 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
 Db 2344 NNN 2403
 Qy 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
 Db 2404 NNN 2463
 Qy 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
 Db 2464 GACTGCAGACTGGAGGAGAACCTGACCCAAAAATATTGCTAGATAATAATATATGACTAT 2523
 Qy 861 GluAsnGlyPheAlaIleGlyProGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
 Db 2524 GAAAAATGGCTTTGCATTTGGACCAAGGTGGCTGGGGTGCAGCTAATAGGCTGATTTACTCT 2583
 Qy 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
 Db 2584 TACATGACTTCTTGGACACTGTGCAGAAACACAGCCACAGCATCGGCATGCCAGTCC 2643
 Qy 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
 Db 2644 TCACGATTAAGAGAGTGGCCCATTTATCTGACTATAAAATTAAGTTAATTTTAAACATC 2703
 Qy 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGln 940
 Db 2704 ACAGCTAGTGTGCCATTTCCCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAA 2763
 Qy 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
 Db 2764 CGACTCTCTTCAGCATTTGGAAACTATCACAATTAATGAAGAGGACTCTCAACAAAGAC 2823
 Qy 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
 Db 2824 CCCATGTATTCCTTTTACGCTTGCAATACAGAAATCTTATAGCCGACAGCAATTCATTAGAA 2883
 Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
 Db 2884 ACAAAAAGGCTTCCCCCTTTCTGCAGACCCAGGCTCAGTGTGCTGAGAGGGCGTATGTGTGTC 2943

```

QY 1001 AenCysProLeuGlyThrTyrTrpAsnLeuGluHisPheThrCysGluSerCysArgile 1020
DB 2944 AATTCGCCCTTTGGGAACCTATTATATCTGGAAACATTTTCACTGTGAAAGCTGCCGGATC 3003
QY 1021 GlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
DB 3004 GGATCTATCAAGATGAGAGGCGCACTTGAGTGCAGCTTTGCCCCCTCTGGGATGAC 3063
QY 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
DB 3064 ACGGAATATATCAATTCAGAAACATCTCTGATGTAAGCTCACTGTGTAACAGGCCACC 3123
QY 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
DB 3124 TACTCATACAGTGGACTTGGAGACTTGTGAATCGTGTCCACTGGGCACATTATCAGCCAAA 3183
QY 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
DB 3184 TTTGGTTCGGGAGCTGCTCTCTGTGTCAGAAACACCTCAACTGTGAAAGAGGAGCC 3243
QY 1101 ValAsnIleSerAlaCysGly 1107
DB 3244 GTGAACATTTCTGCATGTGA 3264

RESULT 2
LOCUS AY405592 3009 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM2275 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405592
VERSION AY405592.1 GI:39761566
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 3009)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3009)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..3009
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
gene
<1..>3009
/locus_tag="HCM2275"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3009
Score: 5060.50 Matches: 928
Percent Similarity: 91.81% Conservative: 3
Best Local Similarity: 91.52% Mismatches: 64
Query Match: 49.52% Indels: 19
DB: 29 Gaps: 1

US-09-977-053-6 (1-1842) x AY405592 (1-3009)

```

```

QY 94 ValGlyGluValAsnPheArgSerGluLeuMetPheValArgLysLeuSerAspPhe 113
DB 1 GTGGCGCAAGTCAACTTCCGACGAGCTCATGTTCGTCGCAAGCTGCTGTCGCACTTC 60
QY 114 ProValValProThrAlaThrArgValAlaIleValThrPheSerSerLysAsnTyrVal 133
DB 61 CCGGTGTGGCCACAGCCGCGCGTGGCCATCGTGACCTTCTCGTCCAGAACTACGTG 120
QY 134 ValProArgValAspTyrIleSerThrArgAlaArgGlnHisLysCysAlaLeuLeu 153
DB 121 GTGGCGCGCGTGCATTTACATCTCCACCCGCGCGCGCAGCACAAAGTGCAGCGCTC 180
QY 154 LeuGlnGluIleProAlaIleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPhe 173
DB 181 TTCCAAGAGATCCCTGGCCATCTCTTACCAGGTGGCGGCCTTACACCAAGGCGCCTTC 240
QY 174 GlnGlnAlaAlaGlnIleLeuLeuHisAlaArgGluAsnSerThrLysValValPheLeu 193
DB 241 CAGCAGCGCGCAATTTCTTCTCATGNNANAGAAACTCAACAAAAGTTATATTTC 300
QY 194 IleThrAspGlyTyrSerAsnGlyGlyAspProArgProIleAlaIleSerLeuArgAsp 213
DB 301 ATCACTGATGATATTTCCAATGGGGNNACCTTAGACCAATTCGCGCTCACTGNNAGAT 360
QY 214 SerGlyValGluIlePheThrPheGlyIleTGTGlnGlyAsnIleArgGluLeuAsnAsp 233
DB 361 TCAGNNNTGGAGATCTTCACTTTGGCATATGGCAGGGAACATTCAGAGCTGATATGAC 420
QY 234 MetAlaSerThrProLysGluGluHisCysTyrLeuLeuHisSerPheGluGluPheGlu 253
DB 421 ATGGCTTCCACCCCAAGNNNNNCACTGTACTCTGCTACACAGTTTGTGANGAATTTGAN 480
QY 254 AlaLeuAlaArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAsp 273
DB 481 NCTTTAGCTCGCGGCGCATTCATGANNATCTACCTTCTGGGAGTTTATTTCAGATGAT 540
QY 274 MetValHisCysSerTyrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySer 293
DB 541 ATGGTCCACTGCTCATATCTTTGTGATGAGCGAGGACTGCTGTGACCGAATGGCAGC 600
QY 294 CysLysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly 313
DB 601 TGCAAAATGGGACACACACAGGCCCATTTTGTAGTGATCTGTGAAAAGGGGTATTACGG 660
QY 314 LysGlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySer 333
DB 661 AAAGGTCTGCAGTATGAATGCACAGCTTGCCTCCATCGGGGACATACAAACCTGAAGCTCA 720
QY 334 ProGlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGly 353
DB 721 CCAGAGGATCAGCAGTTGCAATTCATGCTCTGATGAANAATCACACCTCTCCACCTGGA 780
QY 354 SerThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCys 373
DB 781 AGCACATCCCTGGAAGACTGTGTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGT 840
QY 374 GluLeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThr 393
DB 841 GAACTTGTCCACTGCTGCTGCGGCTCGGAGGCTCCGAAATTTGGTTACTTTATCCAAAACACT 900
QY 394 CysAsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuVal 413
DB 901 TGCAACACCACTTCAATGCAGCCTGTGGGTCCGATGTCACTCCCTGGATTTGATCTTGTG 960
QY 414 GlySerSerIleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCys 433
DB 961 GGAAGCAGCATCATCTTATGTCTACCCAAATGGTTGTGGTCCGGTTCCAGAGACTACTGC 1020
QY 434 ArgValArgThrCysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThr 453
DB 1021 AGAGTAAGAAACATGCTCTCATCTCCGCCAGCCGAAACATGGGCCACATCATGCTGTTTACA 1080

```


Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 3288)

Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarivala, A.,

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,

Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,

Adams, M.D. and Cargill, M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering

them based on alignment.

COMMENT

FEATURES

source

1..3288

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>3288

/locus_tag="HMC2275"

gene

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 3288

Score: 4803.50 Matches: 879

Percent Similarity: 84.74% Conservative: 76

Best Local Similarity: 77.99% Mismatches: 139

Query Match: 47.01% Indels: 33

DB: 29 Gaps: 5

US-09-977-053-6 (1-1842) x AY405593 (1-3288)

Qy 1 MetTrrProArgLeuAlaPheCysCysTrrGlyLeuAlaLeuValSerGlyTrrAlaThr 20
 Db 1 ATGTGGTCCGCGCTGGCCCTTTTGTGCTGGCTCTGGCACTGGTGTCTGGCTGGACCAAC 60
 Qy 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
 Db 61 TTCAGCCCGTGGCCCTTGGCTCACTTCAGCTTCGGCTTCCGCTGTTCCCGAGCCCTCTCG 120
 Qy 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
 Db 121 GGGGCTCTGGCAGACTGGCGGTACTCTCCGCTCCAGTGGAGGAGGAGGAGGAGGAGC 180
 Qy 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgValArgLeuArgGluLeuSer 79
 Db 181 AAAGTGAGAGCGCTGGCGCGCGCTTCCGAGCGCGCTGGCGAGCTGGCGGAGCTCAGC 240
 Qy 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAlaAsnPhe 99
 Db 241 GGCAGCCTTGGAGCTCGTCTTCTGGTGAACGAGTGTCTCAGGTGGCGGCGGCGGCGGCTTC 300
 Qy 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAla 119
 Db 301 CTCACAGAGCTCAAGTTCGTGGCGAAGCTGTCTCGAGCTTCCCGCTGGTGTCCAGCGCC 360
 Qy 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
 Db 361 ACGGTGTGGCCATCGTCACCTTCTCATCCAAAGAACCAACGCTGGCGCGCGGTGATTTAC 420
 Qy 140 IleSerThrArgArgAlaGlnHisLysCysAlaLeuLeuLeuGlnGluLeuProAla 159
 Db 421 ATCTCCACAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Qy 160 IleSerTyrArgGlyGlyGlyThrTrrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
 Db 481 ATCAGCTACCGCGGTGGTGGACCTATACCAAGGCGGCGCTTCAGCAAGCGCGGCAATC 540
 Qy 180 LeuLeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTrrSer 199

Db 541 CTTCTGCTACTAGAGAAAACCTCCAAAGTCATATTTCTCATCCCGAGCGCTATTCC 600
 Qy 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
 Db 601 AATGGCGGAGACCCAGACCTATTCGAGCATCGCTTCGGGATTCGGAGTGAGATCTTC 660
 Qy 220 ThrPheGlyIleTrrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
 Db 661 ACGTTTCGGGATTTGGCAGGGGAATATCCCGGAATCGAATGACATGCTTCACCCCGAAG 720
 Qy 240 GluGluHisCysTrrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAla 259
 Db 721 GAAGAACATTTGTACTCTCCACAGTTTGAAGAAATTTAGAGCTTTAGCTCCAGGCGC 780
 Qy 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyr 279
 Db 781 TTGCATGAAGATCTACTTCTCGGAGTTTATCATCAAGAGGATATGCCCGCTCTCTTAT 840
 Qy 280 LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis 299
 Db 841 CTCCTGTGAGGCTGGGAAAGACTGCTGTACAGAAATGGCCAGCTGCAATGTGGGACAC 900
 Qy 300 ThrGlyHisPheGluCysIleCysGluLysGlyTrrTrrGlyGlyGlyLeuGlnTrrGlu 319
 Db 901 ACGGTCATTTGAATGCACTGTGAGAAGGCTATTACGGAAAGCTCTGCAAGCATGAG 960
 Qy 320 CysThrAlaCysProSerGlyThrTrrLysProGluGlySerProGlyGlyIleSerSer 339
 Db 961 TGCACAGCTTGCCTCATCAGGACATATAGCCGGAAGCTTCTCCAGGAGGAATCAGCAC 1020
 Qy 340 CysIleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAsp 359
 Db 1021 TGCATCCCATGCTCAGACGTAAGCCACACTCCCACTGGAGGACACTTCCCTCGAAGAC 1080
 Qy 360 CysValCysArgGluGlyTrrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
 Db 1081 TCGGTGTCCGAGAGGAGATACCAAGAGATCTGGCAGACTGTGAGGTTGTCTCACTGCT 1140
 Qy 380 AlaLeuLysProProGluAsnGlyTrrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
 Db 1141 GCGCTGAAGCTCTCGAAATGTTTATATACAAACACACTTGCAGAAACCACTTCAT 1200
 Qy 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeu 419
 Db 1201 GCGGCTGTGGGCTCCGATGTCGCGGCTTTGACTTGTGGAGAGCAGCATTCATTTG 1260
 Qy 420 CysLeuProAsnGlyLeuTrrSerGlySerGluSerTrrCysArgValArgThrCysPro 439
 Db 1261 TGTCAACCCATGGTTTGTGTTCTGGACAGAAAGCTTCTGCAGAGTGGAGAGCTGCCCC 1320
 Qy 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTrrLys 459
 Db 1321 CACCTCCAGAGCCCAACACAGCCACATCAGCTGCTCCACTCGCGAAATGTCTACAC 1380
 Qy 460 ThrThrCysLeuValAlaCysAspGluGlyTrrArgLeuGluGlySerAspLysLeuThr 479
 Db 1381 ACCCTGTGTTGTTTACCTGCAATGAAGATACAGATTAGAAGGCGAGCATAGGCTTACC 1440
 Qy 480 CysGlnGlyAsnSerGlnTrrAspGlyProGluProArgCysValGluArgHisCysSer 499
 Db 1441 TGTCAAGGAATGCCAGTGGAGTGGCCGAGAGCCCGGTGTGTAGNACGCCATTTGTGCC 1500
 Qy 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
 Db 1501 ACCTTCAGAGACCCCAAGGCGTCACTATTTCTCCACCCAGCTGCGGAGCAGCGCTGCC 1560
 Qy 520 LysPheGlyThrIleCysTrrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
 Db 1561 AGGCTGGGATGACCTGTGAGTAAAGCTGCCCGCAGGATACATTTTATCCGGGGTCAGA 1620
 Qy 540 GluMetLeuArgCysThrThrSerGlyLysTrrAsnValGlyValGlnAlaAlaValCys 559

1621 GAA---GTGAGATGTGCGACATCTGGCAAGTGGAGTGCCAAAGTTTCAGACAGCTGTGTGC 1677
 560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu 579
 1678 AAGAGATGTGGAGGCTCCACAAATCAGCTGTCCAAATGACATTTGAGGCAGAAAGACTGGGGAG 1737
 580 GlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlu 599
 1738 CAGCAGGACTCTGCTAATGTGACCTGGCAGTCCACAGCTAAGACAACTCTGGTGA 1797
 600 LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
 1798 AAGGTGTCAAGTCCAGCTCCAGCCAGCTTTACCCACCTTACCTCTTCCCAATTGGAGAC 1857
 620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
 1858 GTGGCCATCACCTACACGGCAACGACTCATCCGGTAACCAAGCAGCTGCACCTTCTTAC 1917
 640 IleLysValIleAspAlaGluProValIleAspTyrCysArgSerProProVal 659
 1918 ATTAAGGTCAATGATGTGGAAACCGCTGTGATAGTTGGTCCGATCTTCCACCTCCCAATC 1977
 660 GlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSer 679
 1978 CAGTCTGACAGAGGAGCACCTGCAAGCTGGGATGAGCTCAGTTCTCAGACAACTCC 2037
 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
 2038 GGGGCTGAATGGTTCATTACAGCAGTCCACACAGCGCAGATGTTTCTCATGGGGAA 2097
 700 ThrIleValGlnTyrThrAlaThrAspProSerGlyValAsnAsnArgThrCysAspIleHis 719
 2098 ACGGTGGTGTGTACACGACCTGACCTCCCTCAGGCAACACAGACCTGTGACATCCAC 2157
 720 IleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739
 2158 ATTGTCAATAAA----- 2169
 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
 2170 -----GATAGTGTGGAGTTAACTGTAGCTGAGCTGCAAGGAGGCTATGATTTTC 2220
 760 ThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThr 779
 2221 ACAGAGGGTCAACTAGAGAGTACTACTGTGCTTTTGAAGATGTATCTGGAGACCAACA 2280
 780 TyrThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSer 799
 2281 TACTCTACAGATGGCCAGACTGTGCTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2340
 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
 2341 NNN 2400
 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
 2401 NNN 2460
 840 IleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
 2461 ATTGACTGCAGACTGGAGGAC---CTGACCAAAAAAATACTGCATCGAGTATATTAACAAC 2517
 860 TyrGluAsnGlyPheAlaIleGlyProGlyGlyTrpGlyAlaAlaAsnArgLeuAspTyr 879
 2518 TATGAAAATGGCTTTGCAATTTGACACAGGAGGCTGGGGTGCAGGCAACAGCTGGATAT 2577
 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
 2578 TCTACAGTCACTTCTCTGGATGTTGTACAGAAACACCCCGATGGGCAAGGCCGAGA 2637
 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn 919
 2638 TCGTACGGATTAAAGAACTGTCTCCCATTTCTGACCCCAAAATTCAGCTAAATTTTAC 2697

QY 920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGln 939
 Db 2698 ATCAGCTAGCTGCGCACTCCCGAGGAGAAAGAAAGATACCTTGATTTGATTTGAGGATCAG 2757
 QY 940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959
 Db 2758 CAGCGACTCATTAAGACATTTGGAACAATACCAATCCCTGCAAAAGCACCTTTGAATAAA 2817
 QY 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
 Db 2818 GAGCCCATGATTTCTTCCAGCTGCGCTCGGAAACAGTGGTGGCTGACAGCAATTTCCCTC 2877
 QY 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyValArgMetCys 999
 Db 2878 GAAACAGAAAGGCTTTTCTTCTGACAGACAGGCTCTGTGCTGAGGGGGCGCATGTGT 2937
 QY 1000 ValAsnCysProLeuGlyThrTyrAsnLeuGluHisPheThrCysGluSerCysArg 1019
 Db 2938 GTCAACTGCCCTCGGACCTCTTACTCTGTGGAGCATTCACCTGTGAAAGCTGCCCTC 2997
 QY 1020 IleGlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
 Db 2998 ATGGGATCTTACCAAGATGAAGAGGCGAGCTGGAATGCAAGCTCTGTCCCCCAAGGACT 3057
 QY 1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
 Db 3058 CACAGGAATACCTCCATTTCAAGAGCGTCTCTGAATGCAAGCTCAGTGTGAAGCAAGGC 3117
 QY 1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
 Db 3118 ACCTACTCTTCCAGTGGCTGGAGCTGCGAATGGTCTCGCTGGTACTTATCAACCG 3177
 QY 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
 Db 3178 GAATTTGGATCCCGAGCTGCTCTATGCCAGAAACCAACCAACACCGTGAAGAGGA 3237
 QY 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
 Db 3238 GCGCTGGACATCTCTGCTGTGGA-----GGT 3264
 QY 1120 LeuMetProCysHisProCys 1126
 Db 3265 TATCACCAGTGTTTAGATGC 3285
 RESULT 4
 BC030816
 LOCUS
 DEFINITION Homo sapiens, Similar to polydomain protein, clone IMAGE:4730700,
 mRNA.
 ACCESSION BC030816
 VERSION BC030816.1 GI:21411218
 KEYWORDS HTC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3171)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NTH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbe-remail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcdopaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 37 Row: 9 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction, Similarity but not identity to protein
This clone has the following problem: retained intron.

FEATURES

source

Location/Qualifiers
1. 3171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4730700"
/tissue_type="Placenta"
/clone_lib="NIH MGC 79"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3171
Score: 4781.00 Matches: 867
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 46.79% Indels: 0
DB: 11 Gaps: 0

US-09-977-053-6 (1-1842) x BC030816 (1-3171)

QY	1	MetTProArgLeuAlaPheCysTyrGlyLeuAlaLeuValSerGlyTyrAlaThr	20
DB	338	ATGTGGCCCTGCGCTGTCCTTTGCTGGGGTCTGGCGCTCTGTTGGGGCTGGCGACC	397
QY	21	PheGlnGlnMetSerProSerArgAnPheSerPheArgLeuPheProGluThrAlaPro	40
DB	398	TTTCAGCAGATGTCCTGCTGGCGTTCAGCTTCGGCTCTTCCTCCCGAGACCGCGCC	457
QY	41	GlyAlaProGlySerIleProAlaProAlaProAlaProGlyAspGluAlaGlySerArg	60
DB	458	GGGGCCCGCGGGATATCCCGCGCGCGCTCTGCGGACGACGCGCGGGGACGACA	517
QY	61	ValGluArgLeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGluLeuSerGlu	80
DB	518	GTGGAGCGCTGGGCGGCGGCTTCGGCGGACGCTGGCGCTGCTGGGGAGCTCAGCGAG	577
QY	81	ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAlaAsnPheArg	100
DB	578	CGCTGGAGCTGCTCTTCCTGGTGGATGATTCGTCAGCGTGGCGGCGAAGTCAACTTC	637
QY	101	SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr	120
DB	638	AGCGAGCTCATGTTCGTCGCAAGCTGCTGCGACTTCCTCCGCTGGCGGCGCGCAG	697
QY	121	ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle	140
DB	698	CGGTGGCCATCTGACCTTCTCTGTCAGACTTACGTGGTGGCGCGCTCGATTACATC	757
QY	141	SerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnIleProAlaIle	160
DB	758	TCCACCGCGCGCGCGGACGACAGTGGCGCTGCTCTCCAGAGATCCCTGGCCATC	817
QY	161	SerTyrArgGlyGlyThrTyrThrLysGlyAlaPheGlnAlaAlaGlnIleLeu	180
DB	818	TCCTACCGAGTGGCGGACCTACACCAAGGGGCGCTTCAGCAGCGCGCGCAATTCCT	877
QY	181	LeuHisAlaArgGluLeuSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn	200
DB	878	CTTCATGCTAGAGAAACTCAACAAAGATGTATTTCTCATCTGATGGATATTCGAT	937

QY	201	GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr	220
DB	938	GGGGGAGACCCCTAGACCAATTCAGCGTCACTGCGAGATTTCAGGAGTGAGATCTTC	997
QY	221	PheGlyIleTrrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu	240
DB	998	TTTGGCATATGGCAAGGGAAACATTCAGAGCTGAAATGACATGCTTCCACCCCAAGGAG	1057
QY	241	GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu	260
DB	1058	GAGCACTGTTACTGCTACACAGATTTTGAAGAAATTTAGGCTTTAGCTGCGCGGCA	1117
QY	261	HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerThrLeu	280
DB	1118	CATGAAGATCTACTTCTGGGAGTTTATTTCAAGATGATATGGTCCACTGCTCATATCT	1177
QY	281	CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr	300
DB	1178	TGTGATGAAGGCGAGAGCTGCTGTGACCGAATGGGAGCTGCANAATGTGGACACACA	1237
QY	301	GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys	320
DB	1238	GGCCATTTTGGTGCATCTGTCAAAAGGGGTATTACGGAAAGCTCTGCAGTATGAATG	1297
QY	321	ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys	340
DB	1298	ACAGCTTGGCCATCGGGGACATCAAACTGAAGGCTCCACGAGGAATCAGCAGTGC	1357
QY	341	IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys	360
DB	1358	ATTCCATGCTCTGATGAANAATCACTCTCCACTGGAGACATCCCTCTGAAGACTGT	1417
QY	361	ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla	380
DB	1418	GTCTGCAGAGAGGATACAGGCGCATCTGCCAGACCTGTGAATCTGTCCACTGCCCTG	1477
QY	381	LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla	400
DB	1478	CTGAGACCTCCCGAAATGGTTACTTTATCCAAACACTTGCACCAACCACTTCATATG	1537
QY	401	AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCys	420
DB	1538	GCCTGTGGGGTCCGATGTCACTGATTTGATCTTGTGGGAGAGCATCATCTTATGT	1597
QY	421	LeuProAsnGlyLeuTrrPheSerGlySerGluSerTyrCysArgValArgThrCysProHis	440
DB	1598	CTACCCAAATGTTTGTGGTTCGGTTTCAGAGAGCTTCTGCAGAGTAAAGAACATGTCT	1657
QY	441	LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr	460
DB	1658	CTCCGCGAGCGGAAACATGGCCACATCAGCTGTCTACAGGGGAATGTTATATAGACA	1717
QY	461	ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys	480
DB	1718	ACATGTTGGTTCCTGTGATGAAGGGTACAGACTAGAGGCGAGTGAAGCTTACTTGT	1777
QY	481	GlnGlyAsnSerGlnTrrPheSerGlyProGluProArgCysValGluArgHisCysSerThr	500
DB	1778	CAAGGAAACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGGAGCCCACTGTTCACCC	1837
QY	501	PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys	520
DB	1838	TTTCAGATGCCCAAGATGTCATATATCCCCCAACTGTGGCAAGAGCGAGCCAGCCAA	1897
QY	521	PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu	540
DB	1898	TTTGGGCGATCTGCTATGTAAGTTCGCGCAAGGGTTCAATTTATCTGGAGTCAAGAA	1957
QY	541	MetLeuArgCysThrThrSerGlyLysTrrPheAsnValGlyValGlnAlaLysCysLys	560
DB	1958	ATGCTGAGATGTACCACTTCTGGAAATGGAAATGTCGGAGTTCAGGAGCTGTGTGTA	2017

Qy	561	AspValGluAlaProGlnIleAsnCysProIysAspIleGluAlaIysThrLeuGluGln	580
Db	2018	GACGTGAGGCTCTCAATCACTGCTTAAGACATAGAGGCTAAGACTCTGGAACAG	2077
Qy	581	GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaIysAspAsnSerGlyGluLys	600
Db	2078	CAAGATTCTGCCAATGTTACCTGGCAGATTCCAACAGCTTAAGACAACTCTGGTGAAGAAG	2137
Qy	601	ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal	620
Db	2138	GTGTCAGTCCACGCTTCATCCAGCTTTTACCCACCTTACCTTTTCCCAATTTGGAGATGTT	2197
Qy	621	AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle	640
Db	2198	GCTATCGTATACAGCGCACTGACTATCCGGCAACCCAGGCCAGCTGCATTTTCATATC	2257
Qy	641	LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProProValGln	660
Db	2258	AAGGTTATTGATGCAGAACCACTGTCTATAGACTGGTGCAGATCTCCACCTCCCGTCCAG	2317
Qy	661	ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly	680
Db	2318	GTCTCGGAGAGGTACATGCCCGCAAGCTGGGATGAGCTCTAGTCTCTCAGACACTCAGGG	2377
Qy	681	AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr	700
Db	2378	GCTGATTTGTCATTACCGAAGTCTATACAGAGACCTTTTCCCTCAGGGGAGACT	2437
Qy	701	IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle	720
Db	2438	ATAGTACAGTATACAGCCACTGACCCCTCAGCCCAATACAGACATGTGATATCCATATT	2497
Qy	721	ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys	740
Db	2498	GTCTATAAAGGTTCTCCCTCGTAATTCATTCACCTGTAATGGGGATTTTATATGCG	2557
Qy	741	ThrProAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr	760
Db	2558	ACTCCAGATATATCTGGTGCATCTGATACATTACTTCTGGAGGCTATGATTTCACA	2617
Qy	761	GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr	780
Db	2618	GAAGGGTCTACTGACAAAGTATTATTGCTTATGAGATGGCTCTGGAACCAACATAT	2677
Qy	781	ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe	800
Db	2678	ACCACTGAATGCCAGACTGTGCCAAAAACGTTTTCGAAACCAACGAGTTCAGTCCCTTT	2737
Qy	801	GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu	820
Db	2738	GAGATGTTCTCAAGACGCTGCTGTGTGTGATGACAGATCTGTGAGAGAGTTTCTGAA	2797
Qy	821	AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle	840
Db	2798	GCATTTGAGACGACCTCGGAAAAAATGGTCCCATCATTTTGTAGTGATGCAGAGACATT	2857
Qy	841	AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr	860
Db	2858	GACTGCGAGACTGGAGGAGAACTGCACCAAAAAATATTCCTCCTAGATATATATGACTAT	2917
Qy	861	GluAsnGlyPheAlaIleGly	867
Db	2918	GAATAATGCTTTGCAATTTGGT	2938
RESULT 5	AK052699	2736 bp mRNA linear	HTC 20-SEP-2003
LOCUS	AK052699	Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched	
DEFINITION	AK052699	library, clone:D630026K16 product:polydomain protein, full insert	
ACCESSION	AK052699	sequence.	
VERSION	AK052699.1	GI:26342892	
KEYWORDS	HTC; CAP trapper.		

SOURCE ORGANISM

Mus musculus (house mouse)

Mus musculus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Macsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

REFERENCE

The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2736)

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission

REFERENCE

Submitted (15-JUL-2001) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1. 2736

Qy	1388	GlnSerAenProCysAraGlnAlaThrCysValAspGluLeuAsnSerTyrSerCys	1407	Qy	1748	GluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySer	1767
Db	482	CAGTCCAAACCCATGACAGGACCCAGGCTGGTGGATGAATAAACTCATACAGTTGT	541	Db	1562	GAGTGTGAGTCGGCTCGACTGTAGTGAAGCAGCGCTCTGCTGCAACACCAACGATCC	1621
Qy	1408	LysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSerThrGlyPheAsn	1427	Qy	1768	TyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnGlyCysAlaGluProIle	1787
Db	542	AAATGTCCGCGAGGATTTTCAGGCGCCAGGTTGAGACAGAACAGCCCTTCGGGTTTTAAC	601	Db	1622	TACATATGCTGTGACAGCCCCGTCACGCGAGATGGGAAAAAACTGTGCAGAACTGTT	1681
Qy	1428	LeuAspPheGluValSerGlyIleTyrValMetLeuAspGlyMetLeuProSer	1447	Qy	1788	LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal	1807
Db	602	CTGAGTTTGAAGTTTCTGGCATCTACCGGTACGTTCTTCTAGATGGAGTGCTTCCAACA	661	Db	1682	AAATGTAAAGGCTCCAGAAAAATCCAGAGAAATGCCACTCTTTGGGCAAGATTTATAGCGTG	1741
Qy	1448	LeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAspMetAsnTyrGlyThr	1467	Qy	1808	GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle	1827
Db	662	CTCCATGCCATAACCTGTGCTTCTGGATGAAGTCTCTGATGTCAATCAATACAGGAACG	721	Db	1742	GSTGTGGAAGTCACATTTTCTGTGAAGAAGGCGACCCAGCTGGTGGAGTGGAAAAATC	1801
Qy	1468	ProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsn	1487	Qy	1828	ThrCysLeuGluSerGlyGluTyrAsnHisLeuIleProTyrCys	1842
Db	722	CCCATCTCTATGCACTGGAGGGCAACAAGACAACACCTCTCTCTGACTGATTACAAAT	781	Db	1802	ACGTGTTTGGAGTCTGGCGAGTGGGATCACTCAGACCATCTCTGT	1846
Qy	1488	GlyTrpValLeuTyrValAsnGlyArgGluLysIleThrAsnCysProSerValAsnAsp	1507	RESULT 7			
Db	782	GGCTGGGTTCTTATGTGATGAAGAAAGAAAGATCACCACTGCCCTCTGTGAATGAT	841	AK035333	3242 bp	mRNA	linear
Qy	1508	GlyArgTrpHisAlaIleThrTrpThrSerAlaAsnGlyIleTrpLysValTyr	1527	LOCUS	AK035333.1	GI:26330611	
Db	842	GGCATTTGGCATCATATATGCAATCATCATGGAACAGCACTGGTGGAGCCTGGAGGGTCTAT	901	DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length		
Qy	1528	IleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeuProIleProGly	1547	KEYWORDS	HTC; CAP trapper.		
Db	902	ATAGACGGGAGCTATCTGACAGTGTGTCTGCGCTCTCTGTGGCAAGCCATACCTCTGT	961	SOURCE	Mus musculus (house mouse)		
Qy	1548	GlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLysGlyGluGlyPheSerProAla	1567	ORGANISM	Mus musculus		
Db	962	GCTGTGCTATTAGTTCTTGGCGACAGGACGACGACGACGACGACGACGACGACGACGAC	1021	REFERENCE	1 Carninci, P., and Hayashizaki, Y.		
Qy	1568	GluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrValLeuSerProGln	1587	AUTHORS	Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.		
Db	1022	GAGTCTTTTGGTGGCTCCATGAAGCCAGCTCAACCTCTGGACTATGCTGTCTCCACAG	1081	TITLE	High-efficiency full-length cDNA cloning		
Qy	1588	GlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAla	1607	JOURNAL	99279253		
Db	1082	CAGGTGAAGTCTGCTGCCAGCTCTCCCGCAGAGGAACCTGTCGGGGAAACGTTGTAGCA	1141	MEDLINE	20499374		
Qy	1608	TrpProAspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysSerIlePhe	1627	PUBMED	11042159		
Db	1142	TGGCTGATTTGCTATCGGGATCACGGGAGGTGAAGTTGATTCACGACGATCTTC	1201	REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,		
Qy	1628	CysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAsp	1647	AUTHORS	Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,		
Db	1202	TGCTCTGATTTGCCATCTTTAGAGGGTCCGTGCTCTCACTGAGACCTGCATCAGGAGAT	1261	TITLE	Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,		
Qy	1648	LeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsn	1667	JOURNAL	Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,		
Db	1262	CGAAAGCCAGGTTCANAGTCAAGTGTGTCTGTGACCCCGGCTTCAGATGTCGGGAT	1321	MEDLINE	Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Wakahiki, M.,		
Qy	1668	ProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArg	1687	PUBMED	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,		
Db	1322	CCGTGTCAGTATTGTTGACCAAGGCGAGTGTGTGACAGCCACTCTCTCAGTGTGAACGC	1381	REFERENCE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.		
Qy	1688	IleSerCysGlyValProProLeuGluAsnGlyPheHisSerAlaAspPheTyr	1707	AUTHORS	RIKEN integrated sequence analysis (RISA) system-384-format		
Db	1382	ATTGCTGTGGGCTCCCTCCACCTTGGAGATGGCTTCTATTACGCGAGACCTCCAT	1441	TITLE	sequencing pipeline with 384 multiplexed sequencer		
Qy	1708	AlaGlySerThrValThrTyrClnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArg	1727	JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
Db	1442	GCTGCGACGACAGTACCTATCAGTGCACACGAGTGTCTACTCTGTGGTGACTCCAGA	1501	MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)		
Qy	1728	MetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAsp	1747	PUBMED	20530913		
Db	1502	ATGTTCTGTACGACCAACGGGAGCTGGAAAGCGCATTTCTCATCTCTGCTCGATTTGAT	1561	REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the		
					FANTOM Consortium.		
					Functional annotation of a full-length mouse cDNA collection		
					Nature 409, 685-690 (2001)		
					5 The FANTOM Consortium and the RIKEN Genome Exploration Research		
					Group Phase I & II Team.		
					Analysis of the mouse transcriptome based on functional annotation		
					of 60,770 full-length cDNAs		
					Nature 420, 563-573 (2002)		
					6 (bases 1 to 3242)		
					Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,		

Pukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

ORIGIN

Alignment Scores:

Argument Scores:					
Pred. No.:	2.88e-158	Length:	3242		
Score:	2023.50	Matches:	370		
Percent Similarity:	90.57%	Conservative:	24		
Best Local Similarity:	85.06%	Mismatches:	40		
Query Match:	19.80%	Indels:	1		
DB:	11	Gaps:	1		

US-09-977-053-6 (1-1842) x AK035333 (1-3242)

1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
 239 ATGTGGTCGCGCTGGCCCTTTTGTGTGGGCTCTGGCCTGGTGTGAGCTGGACCAAC 298
 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluIuThrAlaPro 40


```

106 ATCTCCACAGCGCGCCACCAACACAGTGGCGCTACTCAGCGCGAGATCCCGGCC 965
160 IleSerTyrArgGlyGlyGlyThrTyrThySlyGlyAlaPheGlnAlaAalaGlnIle 179
966 ATCACTACCGCGGTGGTGGCACCCTATACCAAGGGCGCCTTCCAGCAAGCGCGCAATC 1025
180 LeuLeuHisAlaArgGluAanSerThrLysValPheLeuIleThrAspGlyTyrSer 199
1026 CTTGCTCCTCTAGAGAAACTCCACCAAGTCAATTTCTCATCACGCGCGCTATTC 1085
200 AspGlyGlyAspProArgProIleAlaAlaSerLeuArgPheSerGlyValGluIlePhe 219
1086 AATGGCGGAGACCCACAGACTATTGCGAGCATCGCTTCGGGATTTGGAGTGGAGATCTTC 1145
220 ThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIys 239
1146 ACGTTGGGATTTGGCAGGGGAAATATCCGGAACTGAATGACATGGCTTCACCCCGAAG 1205
240 GluGluHisCysTyrLeuLeuHisSerPheGlnGluPheGluAlaLeuAlaArgAla 259
1206 GAAGAACATTTTACCTGCTCCACAGTTTGAAGATTTGAGGCTTTAGCTCGCAGGGCG 1265
260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyr 279
1266 TTGCATGAAGATCTACCTCTCTGGAGTTTATCCAGAGGATATGGCCCGCTGCTCTAT 1325
280 LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis 299
1326 CTCGTGGGCTGGGAAAGACTGCTGTGCAGCAATGGCCAGCTGCAAAATGTGGACACAC 1385
300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrGlyLysGlyLeuGlnTyrGlu 319
1386 ACGGTCATTTGATGATCTGTGAAGAGGCTTTTACGGAAAGTCTGCAGCATGAG 1445
320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
1446 TGCACAGCTTGGCCATCAGGACATATAAGCCGGAAGCTTCTCCAGGAGGAATCAGCAC 1505
340 CysIleProCysProAspGluAanHisThrSerProGlySerThrSerProGluAsp 359
1506 TGCATCCCATGCTCTGACGTAAGCCACACCTCCACCTGGAAGCACTTCCCTGGAAGC 1565
360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeu----- 375
1566 TGGTGTGCCAGAGGATACAGAGATCTGGCCGAGACCTGTGAGGG-TAAGCACCCCTAC 1624
375 ----- 375
1625 TGGTCACCAATGGATCTAGAAAGCTGATCTTGATCTCTGTGGTGACAGGTTGGTGTTG 1684
375 ----- 375
1685 AATCCATTTCTTAGGATGATCTTTAACATCCCTGGCTTTGGTCACTTCTGGGGGTAAG 1744
375 ----- 375
1745 GGGGATGGTTGCTTCTGTTTCATAGCATTTGTGAGCACTAAACACATAAGCATCATGGT 1804
375 ----- 375
1805 TCAGAGACGGGATCTGAGGCCATAAGAGTAATGTGATTGACAGAAATATGATGTAGTACT 1864
375 ----- 375
1865 TAGAAATAAAGAAATGCTACAACTTTAGCAATGAATTAGAAATATCATATATAGCATGTCTG 1924
375 ----- 375
1925 ACTCTGGTAGTCTCAAAATTGACTTAAATTTTTCATAAATATTTTCCATCATGATGAAA 1984
375 ----- 375
```

```

1985 GTAGAAAGTATTTATATAGTATACAGACATGCTGGGAAGCTGTGTTTATTGTCTCTATC 2044
375 ----- 375
2045 AATTAGACTGGTTTTTCTCCTTATTTTGTAGTTTATTTTGGTTTTTGTGTTGTT 2104
375 ----- 375
2105 TTTAGTTTCTTCTGTTGTTTGTAGGGTTTTGTTTATTGTTGTTGTTGTTGTTT 2164
375 ----- 375
2165 TTTGGTTTTTGAGATACATTTTATCTGCATAACCTTTGGCTGTCCAGGAACCTCACTCTGTAG 2224
375 ----- 375
2225 GACAGGCTGGCTCAAACTCAAAATATCTCGCCCTGTGCTCCAGAGTCTCGAATT 2284
375 ----- 375
2285 AAAATGGTGTGCCACCACCTCCAGGCTGGTGTCTTCTTCAATTATTTTCAATGTTATTAATA 2344
375 ----- 375
2345 AGAGAGAAAAATATTTAATTATCTATGGAAGTCAAGAACCAAGGACTAGGGTTATATAGCT 2404
375 ----- 375
2405 TGTGGCAGAGCATGTGCTTAGCATACCTGAGATCTCTGGGACTCAGCACCAAGCTGAGGG 2464
375 ----- 375
2465 GCAGGGGAAGGGTGGGAAATCTTACATTTATCTATGTCAGTATCAAAATAAAATCTCTTTG 2524
375 ----- 375
2525 AAGCTTAAAAAGTAGTTGTAGACTATATTAATTTTAAATGGCTAGTATACATCTTCCAG 2584
375 ----- 375
2585 GAGTCATATGTAAAGGTTCTCCCTCAAAACACTCTTAGAAATCAGATGATCAATATAGGA 2644
375 ----- 375
2645 AAGCTATAGGGAATATATGATGATAAATATAGATGTACATATGTAGACAGACAGACAGAT 2704
375 ----- 375
2705 GCATGCTTTGGTAAATAAACTGAAGCAATGCAATGCGTAGACAAACATCTGACTTTGCT 2764
376 -----ValHisCysProAlaLeuLysProProGluAan 386
2765 TTTTGTCTCTCTCTGTAATGTTAGTTGTCACTGTCTGCTGCTGAGGCTCTGAAAT 2824
387 GlyTyrPheIleGlnAanThrCysAanHisPheAanAlaAlaCysGlyValArgCys 406
2825 GGTTTTTTATACAAAACACTTGCAAAACCCACTTCAATGCGCTGTGGGTCGAGTGT 2884
407 HisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsnGlyLeuTyr 426
2885 CGCCCGGGCTTTGACCTTTGGTGGGAAGCAGCATTCATTTGTGTCAACCCCAATGGTTGTGG 2944
427 SerGlySerGluSerTyrCysArg 434
2945 TCTGGGACAGAAAGCTTCTGCAGA 2968
```

RESULT 9

BX390270

LOCUS

DEFINITION

ACCSSION

VERSION

BX390270 906 bp mRNA linear EST 08-MAY-2003
BX390270 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0D1010V21 5-PRIME, mRNA sequence.
BX390270
BX390270.1 GI:30463533

QY	383	ProProGluAsnGlyTyr-PheIleGlnAenThrCysAsnAsnHisPheAsnAlaLysCys	402
DB	422	CCTCCCGAAATGGTACTTTATCCAAACACTTTCGACACACCTTCATGAGCCTGT	481
QY	403	GlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuPro	422
DB	482	GGGGTCCGATGTCACCCCTGGATTTCATCTTGTGGGAAGCAGCATCATCTTATGCTACCC	541
QY	423	AsnGlyLeuTropSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArg	442
DB	542	AATGGTTTGTGGTCCAGTTTCAGAGAGCTACTGACAGAGTAGAAGACATGCTCTCATCTCCGC	601
QY	443	GlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCys	462
DB	602	CAGCCGAAACATGGCCACATCAGCTGTTCTACAGCGAAATGTTATATAAGACACATGT	661
QY	463	LeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGly	482
DB	662	TTGGTTGCTGTGATGATGAGGTACAGACTAGAGGCGGTGATAAGCTTACTTGTCAAGGA	721
QY	483	AsnSerGlnTropAspGlyProGluProArgCysValGluArgHisCysSerThrPheGln	502
DB	722	AACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGAGCGCCACTGTTCCACTTTCAG	781
QY	503	MetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGly	522
DB	782	ATGCCCAAGATGTCATCATATATCCCCACACTGTGGCAGCAGCCAAATTTGGG	841
QY	523	ThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeu	542
DB	842	ACGATCTGCTATGTAAGTTGCGCCGCAAGGTTCATTTATTTTGGAGTCAAGGATTGCTG	901
QY	543	Arg	543
DB	902	AGA	904
RESULT 10			
BX356254			
LOCUS	BX356254	Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA	1201 bp mRNA linear EST 05-MAY-2003
DEFINITION	clone CS0D1010YG21	5-PRIME, mRNA sequence.	
ACCESSION	BX356254		
VERSION	BX356254.1	GI:30380055	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1201)		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4991.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0D1010AD11QPI&cluster=4991.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0D1010AD11QPI. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1010YG21" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) /notes="1st strand cDNA was primed with a NotI-oligo (dT)		

EST.	Homo sapiens (human)	383	ProProGluAsnGlyTyr-PheIleGlnAenThrCysAsnAsnHisPheAsnAlaLysCys	402
SOURCE	Homo sapiens	422	CCTCCCGAAATGGTACTTTATCCAAACACTTTCGACACACCTTCATGAGCCTGT	481
ORGANISM	Homo sapiens	403	GlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuPro	422
REFERENCE	1 (bases 1 to 906)	482	GGGGTCCGATGTCACCCCTGGATTTCATCTTGTGGGAAGCAGCATCATCTTATGCTACCC	541
AUTHORS	Li, W. B., Gruber, C., Jesse, J. and Polayes, D.	423	AsnGlyLeuTropSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArg	442
TITLE	Full-length cDNA libraries and normalization	542	AATGGTTTGTGGTCCAGTTTCAGAGAGCTACTGACAGAGTAGAAGACATGCTCTCATCTCCGC	601
JOURNAL	Unpublished (2001)	443	GlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCys	462
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4991.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0BAG025ZE09_CS02373_1&cluster=4991.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0BAG025ZE09_CS02373_1. Location/Qualifiers 1..906 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1010YG21" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."	661		
FEATURES	source	463	LeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGly	482
		662	TTGGTTGCTGTGATGATGAGGTACAGACTAGAGGCGGTGATAAGCTTACTTGTCAAGGA	721
		483	AsnSerGlnTropAspGlyProGluProArgCysValGluArgHisCysSerThrPheGln	502
		722	AACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGAGCGCCACTGTTCCACTTTCAG	781
		503	MetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGly	522
		782	ATGCCCAAGATGTCATCATATATCCCCACACTGTGGCAGCAGCCAAATTTGGG	841
		523	ThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeu	542
		842	ACGATCTGCTATGTAAGTTGCGCCGCAAGGTTCATTTATTTTGGAGTCAAGGATTGCTG	901
		543	Arg	543
		902	AGA	904
Alignment Scores:				
Pred. No.:	3.44e-131	Length:	906	
Score:	1690.00	Matches:	294	
Percent Similarity:	98.01%	Conservative:	1	
Best Local Similarity:	97.67%	Mismatches:	6	
Query Match:	16.54%	Indels:	0	
DB:	13	Gaps:	0	
US-09-977-053-6 (1-1842) x BX390270 (1-906)				
QY	243	CysTyrLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeuHisGlu	262	
DB	2	TGTTACTCTGACACAGTTTGAAGAAATTGAGGCTTTAGCTCGCGGCAATGTCATGAA	61	
QY	263	AspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyrLeuCysAsp	282	
DB	62	GATCTACTCTTGGGAGTTTATTCAGATGATATGTCACCTCATCTTTGTGAT	121	
QY	283	GluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHis	302	
DB	122	GAAGCGAGAGCTGCTGTGACCAATGGGAGCTGCAATGTGGACACACAGCCAT	181	
QY	303	PheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThrAla	322	
DB	182	TTTGAGTGCATCTGTGAAAGGGGTATTACGGGAAAGTCTGCAAGTATGAATGCACAGCT	241	
QY	323	CysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCysIlePro	342	
DB	242	TGCCCATCGGGGACATACAAACCTGAAGGCTCACAGGAGGATACAGCTTCATTCCTCA	301	
QY	343	CysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCysValCys	362	
DB	302	TGTCCTGATGAATACACCTCTCCACCTCGGAAGCAGCATCCCTCGAAGCTGTGTCTGC	361	
QY	363	ArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLys	382	
DB	362	AGAGAGGGATACAGGGCATCTGGCCAGACCTGTGAAGTGTGCACTTGTCCACTGCCCTGCCCTGAAG	421	
ORIGIN				
LOCUS	1201 bp	1201 bp	linear	EST 05-MAY-2003
DEFINITION	Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
ACCESSION	clone CS0D1010YG21			
VERSION	BX356254			
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 1201)			
AUTHORS	Li, W. B., Gruber, C., Jesse, J., and Polayes, D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4991.r For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0D1010AD11QPI&cluster=4991.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0D1010AD11QPI. Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1010YG21" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT)			


```

|||||
421 GAAATTCATTACACCTGTAAATGGGATTTTATATGCACTCCAGATATACTGGAGTC 480
|||
748 AATCCTGTTTATGAGTTCCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGGCT 767
|||
481 AACTGTACATTAATCTGCTTGGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGGCT 540
|||
768 TTTCTGCTTATGAGTTCCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGGCT 787
|||
541 TATTTGCTTATGAGTTCCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGGCT 600
|||
788 AATCTGTTTATGAGTTCCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGGCT 807
|||
601 GCAAAACCAAGCTTTTGCACCAACCAAGCTTTTGCACCAACCAAGCTTTTGCACCA 660
|||
808 ATGCTGTTGATGACACGATCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGG 827
|||
661 CGTTGTTGATGACACGATCTGTTGAGGCTGTTGAGGCTGTTGAGGCTGTTGAGG 720
|||
828 LysMetValProSerPheCysSerAsp 836
|||
721 AATATGCTCCATCATTTTGTAGTGAT 747

```

```

RESULT 13
AL706149
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

AL706149 738 bp mRNA linear EST 04-SEP-2003
 DKFZp686E243 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DKFZp686E243 5', mRNA sequence.
 AL706149
 AL706149.1 GI:19689504
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and
 Wiemann, S.
 EST (Duesterhoeft, et al.)
 Unpublished (1999)
 Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No 61 sequence available.
 This clone (DKFZp686E243) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

```

Source
Location/Qualifiers
1..738
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686E243"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.69e-103 Length: 738
Score: 1359.00 Matches: 238
Percent Similarity: 98.36% Conservative: 2
Best Local Similarity: 97.54% Mismatches: 4
Query Match: 13.30% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-977-053-6 (1-1842) x AL706149 (1-738)
QY 1253 PheileCysGluCysProSerGlyThrGlyGlnArgCysGluGluAsnIleAsnGlu 1272
DB 3 TTTATTTGAGTTCCTGTTGAGTTCCTGTTGAGTTCCTGTTGAGTTCCTGTTGAGTTC 62
QY 1273 CysSerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArg 1292
DB 63 TGTAGCTCCAGTCTCTGTTTAAATAAAGGAATCTGTTGATGTTGTTGCTGGCTATCGT 122
QY 1293 CysThrCysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGln 1312
DB 123 TGCACATGTTGTGAAGGATTTGTAGGCTGTCATTTGTGAAGCAGAAATCNMTGAATGCCAG 182
QY 1313 SerAsnProCysLeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuValAsp 1332
DB 183 TCAACCATGCTTAAATAATGCACTCTGTGAAGACCAAGTTGGGGGATTTCTTGTGCA 242
QY 1333 CysProProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGln 1352
DB 243 TGCCCACTCGATTTTGGGTACCGATGTGGAAGAAGCTGCGATGAGTGTCTCAGTCAG 302
QY 1353 ProCysLysAsnGlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuVal 1372
DB 303 CCATGCAAAATGGAGCTACCTGTAAAGACGTCCTCAATAGCTTTCAGATGCTGTGTGCA 362
QY 1373 AlaGlyPheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCys 1392
DB 363 GCTGGCTTTCACGATCACCTGTGAATTAACATCAATGAATGTCAATGCTCAATCCATG 422
QY 1393 ArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGly 1412
DB 423 AGAATCAGCCACATCTGTGTGATTAATTAATTCATACAGTTGTAATGTGAGCCAGGA 482
QY 1413 PheSerGlyGlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluVal 1432
DB 483 TTTTCAGGCAAAAGGTGTGAACAGACAGCTACAGGCTTTAACTGATTTTGAAGTT 542
QY 1433 SerGlyIleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThr 1452
DB 543 TCTGGTATCTATGATGATGCTAGTGGCTCTCCATCTCTCATGCTCTTAACC 602
QY 1453 CysThrPheTrpMetLysSerSerAspMetAsnTyrGlyThrProIleSerTyrAla 1472
DB 603 TGTACCTTCTGATGAATCTCTGACGACATGAATGATGACACCAATCTCTATGCA 662
QY 1473 ValAspAsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyr 1492
DB 663 GTTGATAACGGCAGCAGCAATACCTCTCTGACTGATTATAACGGCTGGGTTCTTTAT 722
QY 1493 ValAsnGlyArg 1496
DB 723 GTGAATGGCAGG 734

```

```

RESULT 14
AL706149
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

AL706149 801 bp mRNA linear EST 02-AUG-2002
 AL706149 PLACE1 Homo sapiens cDNA clone PLACE1006157 5', mRNA
 sequence.
 AL706149
 AL706149.1 GI:10997819
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
 Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
 Isogai, T.
 HRI human cDNA project
 Unpublished (2000)
 Contact: Takao Isogai

Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix
 Research Institute; cDNA library construction: Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.

FEATURES
 source
 1..801
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1006157"
 /tissue_type="placenta"
 /clone_lib="PLACE1"
 /note="Vector: pMB18SFL3"

ORIGIN
 Alignment Scores:
 Pred. No.: 8,31e-99 Length: 801
 Score: 1304.00 Matches: 247
 Percent Similarity: 91.18% Conservative: 1
 Best Local Similarity: 90.81% Mismatches: 16
 Query Match: 12.76% Indels: 8
 DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x AUI37280 (1-801)

QY 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyLeu 337
 DB 1 TATGAATGACACAGCTTGCCCATCGGGACATACAAACCTGAAGCCTCACCAGGAGGAATC 60
 QY 338 SerSerCysLeuProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
 DB 61 AGCAGTGTGATCTCCATGTCCTCGATGAATATCACACCTCTCCACCTGGAAGCACATCCCT 120
 QY 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
 DB 121 GAAGACTGTCTGCAGAGAGGATACAGGCATCTGGCAGACCTGTGACTGTGCCAC 180
 QY 378 CysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHis 397
 DB 181 TGCCCTGCTCGAAGCTCCCGAAATGTTACTTTATCCAAACACCTTGCACCAACCCAC 240
 QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAsnLeuValGlySerSerIle 417
 DB 241 TTCATGTGACGCTGTGGGGTCCGATGTACCCCTGGATTTGATCTTGTGGAGACGACATC 300
 QY 418 IleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThr 437
 DB 301 AUCTTATGTCTACCCAAATGTTTGTGTGTCGGTTTCAGAGAGTACTTCAGAGTAAGAACA 360
 QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
 DB 361 TGTCTCTCATCTCGCCAGCGCAACATGCCACATCAGCTGTGTTCTACAGGGAAATGTTA 420
 QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
 DB 421 TATAAGACAACATGTTTGTGTTGCTGTGTATGAAGGGGTACAGACTAGAGGACGTGATTAAG 480
 QY 478 LeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHis 497
 DB 481 CTTACTTGTTCAGGAACACGCACTGGATGGCCAGAACCCCGGTGTGTGGAGGCCAC 540
 QY 498 CysSerThrPheGlnMet-ProLysAspValIleIleSerProHisAsnCys-GlyLysG 517
 DB 541 TGTTCACCTTTTCAGATGCCCAGAAAGATGTTCATCATATCCCCCACAACCTGTGGGCAAGC 600
 QY 517 InProAlaLysPheGlyThrIleCysTyrVal-SerCys-ArgGlnGlyPheIleLeuSe 536

601 ANCCAGCCAAATTTGGACGATCTGCTATGTTAAGTTGCCCGCAAGGGTTCAATTTATC 660
 QY 536 rGlyValLysGluMetLeuArgCysThrThrSerGlyLys--TrpAsnValGlyValGln 555
 DB 661 TGGAGTCAAGAAATGCTGATATGTACAACTCTTCGGGAAATTTGGAATGTCGAGTTCAA 720
 QY 556 -AlaAlaVal-CysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluA 575
 DB 721 GGAGCTGTGTGTAAAGACNTGGAAGGNCCTCATCACTGCTCTTAGGNCATTANAG 780
 QY 575 lAlLysThrLeuGluGlnGln 591
 DB 781 GGTAAAGACTCNGGGAACAG 800

RESULT 15
 BG620159 756 bp mRNA linear EST 18-APR-2001
 LOCUS 602618422F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4732181 5',
 DEFINITION mRNA sequence.
 ACCESSION BG620159 GI:13671530
 VERSION BG620159.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 756)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI591 row: 1 column: 06
 High quality sequence stop: 738.

FEATURES
 source
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4732181"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 79"
 /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggccgcttggcc); Site_2: SfiI
 (ggccattatggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CAGGCCATTATGGCC-3'
 and 3' adaptor sequence:
 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.3
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 4,29e-98 Length: 756
 Score: 1295.00 Matches: 250
 Percent Similarity: 98.04% Conservative: 0
 Best Local Similarity: 98.04% Mismatches: 1
 Query Match: 12.67% Indels: 5
 DB: 12 Gaps: 0

US-09-977-053-6 (1-1842) x BG620159 (1-756)

QY 476 AspLysLeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGlu 495

1 GATAAGCTTACTTGTCAAGGAAACAGCCAGTGGATGGCCAGAACCCCGGTGTGTGGAG 60
 496 ArgHisCysSerThrPheGlnMetProLysAspValIleLeuSerProHisAsnGlyGly 515
 61 CGCCACTGTGTCACCTTTTCAGATGCCCAAGATGTCATCATATCCCCCAACACTGTGGC 120
 516 LysGlnProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeu 535
 121 AAGCAGCCAGCCCAATTTGGGACGATCTGCTATGTAAGTGGCCCAAGGGTTCATTTTA 180
 536 SerGlyValLysGlnMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGln 555
 181 TCTGGAGTCAAGAAGAAATGCTCAGATGTACCACTTCTCGAATAATGAATGTGCGAGTTCAG 240
 556 AlaAlaValCysLysAspValGluAlaProGlnIleLeuSerProLysAspIleGluAla 575
 241 GCAGCTGTGTAAAGAGCTGGAGGCTCTCAATCAACTGTCTTAAGGACATAGAGGCT 300
 576 LysThrLeuGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAsp 595
 301 AAGACTTGGACAGCAAGATTCGCCATGTCTACCTGGCAGATTCACACAGCTAAAGAC 360
 596 AsnSerGlyGluLysValSerHisValHisProAlaPheThrProProTyrLeuPhe 615
 361 AACTCTGTGTAAGAGGTGTCAGTCCAGCTTCATCCAGCTTTCACCCACCTTACCTTTTC 420
 616 ProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSer 635
 421 CCAATTGGAGATGTGTCTATCGTATACACGGCAACTGACCTATCCGGCAACAGCCACG 480
 636 CysIlePheHisIleLysValIleAspAlaGluProProValIleAspTyrCysArgSer 655
 481 TGCATTTTCCATATCAAGTTATTGTATGTCAGAACCACTGTCTATGACTGTGTGAGATCT 540
 656 ProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPhe 675
 541 CCACCTCCCGTCCAGGTCCTCGAAGAGGTACATCGCGCAAGCTGGGATGAGCCCTCAGTTC 600
 676 SerAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPhe 695
 601 TCAGACAACCTCAGGGCTGGAATTTGTCATTACCAAGAGTATACACCAAGGAGACCTTTTC 660
 696 ProGln-GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgTh 715
 661 CCTCAAGGGGAGACTATGATGATGATATACAGTATACAGC-ACTGACCCCTCAGGCATACAGGAC 719
 715 r-CysAspIle-HisIleVal-IleLysGlySerPro 726
 720 CATGTGATATCCCATATTGTCCATAAAGGTTCTCCG 756

RESULT 16

EG434540 829 bp mRNA linear EST 14-MAR-2001
 LOCUS 602506592P1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4603823 5',
 DEFINITION mRNA sequence.

ACCESSION EG434540.1 GI:13341046

VERSION EST.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 829)

NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgaps@emil.nih.gov

Tissue Procurement: CLONTECH Laboratories, Inc.

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCM1338 row: g column: 24

High quality sequence stop: 676.

Location/Qualifiers

source

1. 829

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4603823"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: placenta; Vector: pDNR-LIB (Clontech);

Site 1: SfiI (ggcgctcgcc); Site 2: SfiI

(ggcattatggcc); 5' and 3' adaptors were used in cloning

as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'

and 3' adaptor sequence:

5'-ATTCTAGAGCGGAGCGCGCCGACATG-dT(30)BN-3' (where B = A,

C, or G and N = A, C, G, or T). Average insert size 1.3

kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

by PCR. This library was enriched for full-length clones

and was constructed by Clontech Laboratories (Palo Alto,

CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 5,08e-89 Length: 829
 Score: 1188.00 Matches: 240
 Percent Similarity: 94.14% Conservative: 1
 Best Local Similarity: 93.75% Mismatches: 5
 Query Match: 11.63% Indels: 11
 DB: 12 Gaps: 2

US-09-977-053-6 (1-1842) x BG434540 (1-829)

Qy 618 GlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
 Db 1 GGAGATGTTGCTATCGTATACACGGCAACTGACCTATCCGGCAACAGGCCAGCTGCAIT 60
 Qy 638 PheHisIleLysValIleAspAlaGluProProValIleAspTyrCysArgSerProPro 657
 Db 61 TTCATATCAAGGTTATTGTATGTCAGAACCCACTGTCTAGACTGTGTGCAAGATCTCCACCT 120
 Qy 658 ProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAsp 677
 Db 121 CCCTCCAGGTCCTCGAGAGGTATCATCCGCAAGCTGGGATGAGCCTCAGTCTCTCAGAC 180
 Qy 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697
 Db 181 AACTCAGGGGCTGAATTTGTCATTACCAAGAGTCTATACACAGGAGACCTTTTCCCTCAA 240
 Qy 698 GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717
 Db 241 GGGGAGACTATAGTACAGTATACAGCCACTGACCCCTCAGGCCAATATACAGGACATGTGAT 300
 Qy 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
 Db 301 ATCCATATTGTTCATATAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGAT 360
 Qy 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
 Db 361 TTTATATGCACTCCAGATAATTAAGTGGAGTCAACTGTACATTAACCTTGTGAGGGCTAT 420
 Qy 758 AspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLys 777
 Db 421 GATTTCACAGAGGGTCT---GACAAGTATTATTGTGTCTTATGAAGATGGCGCTCTGGAAA 477
 Qy 778 ProThrTyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPhe 797
 Db 478 CCAACATATACACTGAATGCCAGACTGTGCCAAAAAAGC-TTTTGAAACACCGGGTTC 536
 Qy 798 LysSerPheGluMet-PhetylLysAlaAlaArgCysAspThrAspLeuMethylsly 817

Db 537 AAGTCTTTGAGATGGTCTTACAAAGCAGCTGTTGTGATGACACAGATCTGATGAAGAA 596
 QY 817 sPheSerGluAlaPheGluThrThrLeuGly-LysMetValProSerPheCysSerAspa 837
 Db 597 GTTTCTGAAGCATTTGAGACGACCTGGCAAAAATGGTCCCATCAATTTGTAGTGATG 656
 QY 837 laGlu-AspIleAspCysArgLeuGlu---GluAsnLeuThrLysLysTyr-CysLeuGI 855
 Db 657 CAGAGGACATTTGACTGCGACGTGAAGAGAGAAACCTGAGCAAAAAAATTTTGTCTTGA 716
 QY 855 u-TyrAsn-TyrAsp-TyrGlu-AsnGlyPhe 864
 Db 717 AATATAATTTATGACTTATGAATAATGGGTTT 748

RESULT 17
 LOCUS BU929100 701 bp mRNA linear EST 18-OCT-2002
 DEFINITION AGENCOURT_10425850 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:5662868
 5', mRNA sequence.
 ACCESSION BU929100
 VERSION BU929100.1 GI:24117830
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 701)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9abps-remail.nih.gov
 Tissue Procurement: CLONTECH
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLCM2928 row: a column: 12
 High quality sequence stop: 627.

FEATURES

Location/Qualifiers
 1..701
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5662868"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 78"
 /note="Organ: Placenta; Vector: pDNR-LTB (Clontech);
 Site_1: SfiI (ggccattatggcc); Site_2: SfiI
 (ggccattatggcc); 5' and 3' adaptors were used in cloning
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'
 and 3' adaptor sequence:
 5'-ATTCTAGACCGGCGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.3
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3,5e-88 Length: 701
 Score: 1176.50 Matches: 219
 Percent Similarity: 99.10% Conservative: 0
 Best Local Similarity: 99.10% Mismatches: 1
 Query Match: 11.51% Indels: 2
 DB: 13 Gaps: 1
 US-09-977-053-6 (1-1842) x BU929100 (1-701)

QY 737 AspPheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGly 756
 Db 3 GATTTTATATGACCTCCAGATATATCTGGAGTCACTGATACATTAATCTGCTTGGAGGCG 62
 QY 757 TyrAspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyr 776
 Db 63 TATGATTTACAGAGGGTCTACTGCAAGTATTAATTTGTCTTATGAAGATGGCGTCTGG 122
 QY 777 LysProThrTyrThrThrGluTTPProAspCysAlaLysLysArgPheAlaAsnHisGly 796
 Db 123 AAACCAACATATACCACTGAATGGCAGACTGTGGCCAAAAAAGCTTTTGCACCAACCGG 182
 QY 797 PheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLys 816
 Db 183 TTCAGTCTCTTGAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAG 242
 QY 817 LysPheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAsp 836
 Db 243 AAGTTTCTGAAAGCATTTGAGACGACCTGGGAAAAATGGTCCCATCATTTTGTAGTGAT 302
 QY 837 AlaGluAspIleAspCysArgLeuGluAsnLeuThrLysLysTyrCysLeuGluTyr 856
 Db 303 GCAGAGGACATTTGACTGCGAGCTGGAGGAGAACCTGACCAAAAAAATATTCCTAGAATAT 362
 QY 857 AsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAsnArg 876
 Db 363 AATTATGACTATGAATAATGGCTTTCATTTGGACAGGTGGCTGGCGTGCAGCTAATAGG 422
 QY 877 LeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGly 896
 Db 423 CTGGATTACTCTTACGATGACTCTCTCGACACTGTGCAGAAACAGCCACCAAGCATCGGC 482
 QY 897 AsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeu 916
 Db 483 AATGCCAAGTCTCTCAGGATTAAGAAGTGCCCATTTATCTGACTATAAATAATTAAGTTA 542
 QY 917 IlePheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyr 936
 Db 543 AATTTTAACTACACAGT----GTGCCATTACCCGATGAAGAAGATGATACCTTTGAATGG 598
 QY 937 GluAsnGlnGlnArgLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThr 956
 Db 599 GAAATATCAGAACGACTCTCTTCAGACATTTGGAAACTATCACAATAAATACTGAAAGGACT 658
 QY 957 Leu 957
 Db 659 CTC 661

RESULT 18

LOCUS BG436029 719 bp mRNA linear EST 14-MAR-2001
 DEFINITION 602508809P1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4619402 5',
 mRNA sequence.
 ACCESSION BG436029
 VERSION BG436029.1 GI:13342535
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 719)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: c9abps-remail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>

Plate: LLC1379 row: a column: 03
 High quality sequence stop: 696.

FEATURES
 source
 1..719
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4619402"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_79"
 /note="Organ: placenta; vector: pNMR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGCGCCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.64e-85. Length: 719
 Score: 1145.00 Matches: 224
 Percent Similarity: 96.58% Conservative: 2
 Best Local Similarity: 95.73% Mismatches: 8
 Query Match: 11.20% Indels: 6
 DB: 12 Gaps: 0

US-09-977-053-6 (1-1842) x BG436029 (1-719)

Qy 630 SerGlyAsnGlnAlaSerCysPheHisIleValIleAspAlaGluProVal 649
 Db 2 TCCGGCAACCAAGGCGAGCTGCAATTTTCCATATCAAGGTTATTGTCAGAACCACTGTC 61
 Qy 650 IleAspTrpCysArgSerProProValGlnValSerGluValHisAlaAlaSer 669
 Db 62 ATACGTGTGTCAGATCTCCACCTCCCGTCAGGTCCTCGAGAGGTACATGCCCAAGC 121
 Qy 670 TrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHis 689
 Db 122 TGGGATGAGCTTCAGTTCTCAGACCACTCAGGGCTGAATGGTCAATACCAAGAGTCAT 181
 Qy 690 ThrGlnGlyAspLeuPheProGlnGlyThrIleValGlnThrAlaThrAspPro 709
 Db 182 ACACAAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAGTATACAGCCACTGACCCC 241
 Qy 710 SerGlyAsnAsnArgThrCysAspIleHisIleValIleValGlySerProCysGluIle 729
 Db 242 TCAGGCAATAACAGACATGATGATATCCATATGTCATAAAGGTTCTCCCTGTGAATTT 301
 Qy 730 ProPheThrProValAsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCys 749
 Db 302 CCATTACACCTGTAAATGGGGAATTTATATGCACTCCAGATATATCTGGAGTCAACTGT 361
 Qy 750 ThrLeuThrCysLeuGluGlyThrAspPheThrGluGlySerThrAspLysThrCys 769
 Db 362 ACATTAACTTCTTGGAGGGCTATGATTCACAGAAAGGCTCTACTGCAAGTATTATTGT 421
 Qy 770 AlaThrGluAspGlyValTrpLysProThrThrThrThrThrThrThrThrThrThrThr 789
 Db 422 GCTTATGAAGATGGCGTCTGGAACCAACATATACCTGAATGCGCAGATGTGCCAANA 481
 Qy 790 LysArgPheAlaAsnHisGlyPheLysSerPheGluMetPheThrLysAlaAlaArgCys 809
 Db 482 AAACGTTTGCACCAACCGGGTCTCAAGTCTCTTGAAGATGTTCTCAAAAGCAGCTGTTGT 541
 Qy 810 AspAspThrAspLeuMetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMet 829
 Db 542 GATCACACAGATCTGATGAAGAGTGTCTTGAAGCATTTTGACGACCTG-GGAAAAATG 600

Qy 830 ValProSerPheCysSerAspAlaGluAspIleAspCysArgLeuGluGluAsnLeuThr 849
 Db 601 GTCCCATCATTT-TGTAGTGTATGTCAGAGGACATTTGATGG-AAGATGGAGGAACTG-AAC 657
 Qy 850 LysLysThrCysLeuGluThrAsnThrAspThrGluAsnGly 863
 Db 658 AAAAATATGG-CTAGAATATAATATTATGA-TATGAAATGGT 697

RESULT 19
 AK045217 1887 bp mRNA linear HTC 20-SEP-2003
 LOCUS Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length
 DEFINITION enriched library, clone: B130048F05 product: polydomain protein, full
 insert sequence.
 ACCESSION AK045217
 VERSION AK045217.1 GI:26090780
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 3 Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
 Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multipillar sequencer
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE 20530913
 PUBMED 11076861
 4 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 5 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL Nature 420, 563-573 (2002)
 6 (bases 1 to 1887)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Tgami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. Tomohiro Kono (Department of Animal Science, Tokyo University of Agriculture, 1737 Hunko Atsugi City, Kanagawa Prefecture, Japan) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

1. 1887
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:B130048F05"
/db_xref="MGI:2410576"
/db_xref="taxon:10090"
/clone="B130048F05"
/tissue types="parthenogenote"
/clone lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="9.5 days embryo"
1. 1887
/notes="polydomain protein (MGI:MGI:1928849, GB|NM_022814, evidence: BLASTN, 100%, match=1346)"

misc_feature

ORIGIN

Alignment Scores:
Pred. No.: 3,97e-83 Length: 1887
Score: 1125.50 Matches: 220
Percent Similarity: 88.97% Conservative: 14
Best Local Similarity: 83.65% Mismatches: 28
Query Match: 11.01% Indels: 1
DB: 11 Gaps: 1

US-09-977-053-6 (1-1842) x AK045217 (1-1887)

QY 1 MetTTPProArgLeuAlaPheCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr 20
Db 314 ATGTGGTGGCGCTGGCTTTTGTGTGGCTCTGGCACTGGTGTGAGTGGACCAAC 373
QY 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
Db 374 TTCAGGCCGTGGCCCTTCGCTCACTTCAGCTTCGGCTGTTCCTCCGAGGCTCTCCG 433
QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
Db 434 GGGGCTCTGGCGAGCTGGCGGTACCTCCCGCTCCAGTGGAGGAGGAGCGAGGAGC 493
QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSer 79
Db 494 AAAGTGGAGCGCTGGCGCGCGCTTCGGAGCGCGCTGGCGGAGCTGGCGAGCTCAGC 553
QY 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerSerValGlyGluValAsnPhe 99
Db 554 GGCAGCGCTGGAGCTCGCTTCCTCGTGGAGCGAGCTGCTCCAGCGGTGGCGCAACCACTTC 613
QY 100 ArgSerGluLeuMetPheValArgGlyLeuLeuSerAspPheProValValProThrAla 119
Db 614 CTCAACGAGCTCAAGTTCGTGGCGGAGCTGTGCTCCGACTTCCCGGTGGTGTCCAGCGCC 673
QY 120 ThrArgValAlaIleValThrPheSerSerIysAsnTyrValValProArgValAspTyr 139
Db 674 ACGGTGTGGCCATCGTCACCTTCTCATCCAGAACCAACGTTGGTGGCGCGGTGATTAC 733

QY 140 ILeSerThrArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluLeuProAla 159
Db 734 ATCTCCACCAAGCGCGCGCCCAACCAAGTGGCGCTACTCAGCCGCGAGATCCCGGCC 793
QY 160 ILeSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
Db 794 ATCACTACCGCGTGGTGGCACTATACCAAGGGCGCTTCCAGCAAGCGCGCAATC 853
QY 180 LeuLeuHisAlaArgGluLeuSerThrLysValValPheLeuLeuThrAspGlyTyrSer 199
Db 854 CTTGCTCACTTAGAGAAACTCCACCAAGTCATATTTCTCATCACGCGCTATTTC 913
QY 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgPheSerGlyValGluLeuPhe 219
Db 914 AATGGCGGAGACCCCGACCTATTTCAGCATTCGCTTCGGGATTTCCGAGTGGAGATCTTC 973
QY 220 ThrPheGlyIleTTPGlnGlyAsnIleArgGluLeuLeuAsnMetAlaSerThrProLys 239
Db 974 ACCTTCGGGATTTGGCAGGGGGAATATCCGGGAAGTGAATGACATGGCTTCCACCCGGAAG 1033
QY 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
Db 1034 GAAGACATTTGTTACCTGCTCCACAGTTTGAAGATTTGAGGCTTTAGTCTCGAGGCG 1093
QY 260 LeuHisGlu 262
Db 1094 TTGCATGAA 1102
RESULT 20
AK087058
LOCUS
DEFINITION
Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched library, clone:E030024D18 product:polydomain protein, full insert sequence.
ACCESSION
VERSION AK087058.1 GI:26103990
KEYWORDS
SOURCE HTC; CAP trapper.
ORGANISM
Mus musculus (house mouse)
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
AUTHORS
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
AUTHORS
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
AUTHORS
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
AUTHORS
TITLE Functional annotation of a full-length mouse cDNA collection

/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',
 GAGAGAGAGCGCGCGCACTCGAGTTTCTTTTCTTTTCTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5',
 GAGAGAGAGATTCGAGTTAATTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I."

ORIGIN

Alignment Scores:
 Pred. No.: 1.6e-72 Length: 673
 Score: 990.00 Matches: 175
 Percent Similarity: 86.61% Conservative: 19
 Best Local Similarity: 78.12% Mismatches: 28
 Query Match: 9.63% Indels: 2
 DB: 10 Gaps: 0

US-09-977-053-6 (1-1842) x BB542219 (1-673)

QY 1045 HisSerArgAsnIleSer-AspCysLysAlaGlnCysLysGln-GlyThrTyrSerTyrS 1064
 DB 6 CATTCAAGAGGGTTCTTAATGCAAGCTCAGTGTAAGCAGGGCCCTTACTCTTCCA 65
 QY 1064 erGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySerA 1084
 DB 66 GTGGGTGGAGACCTGCGAATCGTTCGTGGGTACTTATCAACGGGAATTTGATGCC 125
 QY 1084 xgSerCysLeuSerCysProGluSerThrSerThrValIysArgGlyAlaValAsnIleS 1104
 DB 126 GGAGCTGCTCTCTAATGCGCAACCAACCAACCAACGGTGAAAGAGGCGGTGACATCT 185
 QY 1104 erAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysH 1124
 DB 186 CTGCTTGTGGAGTGCCTGCGCCAGTAGGAGAAATCTCCGCTTCTGGGTACACACCTGCT 245
 QY 1124 isProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysP 1144
 DB 246 ACCCTTGCCCTCGAGACTATTACCAACCAATGCAAGGAGAGTCTCTGCTCGCTGCTGC 305
 QY 1144 roPheTyrGlyThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheS 1164
 DB 306 CCTTTTATGGAACTACCAACCATCTGCGCGCCACGTCCTCATCAGAGCTGCTCAAGTTTA 365
 QY 1164 erSerThrPheSerAlaAlaGluSerValValProAlaSerLeuGlyHisIleL 1184
 DB 366 GCTCTACTTCTCAGCAGCAGAAAGACATAGTGCCCTCTGTCGCCCTCGACATTCCTCC 425
 QY 1184 ysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisA 1204
 DB 426 AGAACAAGTACGAAGTCAGCAGTCTTTTCAAGAAATGCTTCTTAAACCCCTGCCACA 485
 QY 1204 snSerGlyThrCysGlnGluLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrT 1224
 DB 486 ACAGTGAACCTGCAACAGCTTGGGGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 545
 QY 1224 hrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnG 1244
 DB 546 CAGGCTTAAGTGTGAACAGATATTGATGAATGCAGCTCTCTGCTTGTCTGCTCAATGCTG 605
 QY 1244 lyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyG 1264
 DB 606 GAATTTGTAGACCAAGTTGGGGGATTCAGTGGCAATGTTCATTTGGGCTATTTCAGGTC 665
 QY 1264 lnArgCys 1266

Db 666 AAATATGT 673

RESULT 25

BQ717243

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..998

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6196142"

/sex="male"

/tissue_type="sympathetic trunk"

/dev_stage="adult, 16 yr"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:

NotI; Site 2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACCCAGCGTCCG-3' and

5'-GACATGTTCTAGATCGAGCGCGCCCT(15)-3'. Size selected

1 kb for average insert length 1.9 kb. This is a primary

library, non-amplified. Library constructed by Life

Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor

College of Medicine); available through Life

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

```

QY 312 TyrGlyLysGlyLeuGlnTyrGluCys----- 320
Db 181 TACGGAAAGGCTGCAGTATGAATGCACAGCTGATGTGCCACATAACATGGACTTCTCC 240
QY 321 -----ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyG1 336
Db 241 GGTCTTCTGCCCCACAGCTTCCCATCGGGGACATACAACTGAGGCTCACCAGGAG 300
QY 336 YleSerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSe 356
Db 301 AATCAGCAGTTCATTCCATGCTCATGAAATCACACCTCTCCACCTGGAAGCACATC 360
QY 356 rProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuVa 376
Db 361 CCGTGAAGACTGTGTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGT 420
QY 376 lHisCysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAs 396
Db 421 CCACTGCCCTGCCCCGAGCTCCGAAATGGTTACTTTATCCAAACACCTTGGACAA 480
QY 396 nHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySer-S 416
Db 481 CCACTTCAATGCAGCCTGTGGGTCCGATGTCACTCCCTGGATTGATCTTGTGGGAGCAA 540
QY 416 erIleIleLeuCysLeuProAsnGlyLeuTrpSer-GlySerGluSerTyrCysArg 434
Db 541 GCATCCTCTTATGTCTACCCCAATGGTTGTGGGCCCGGTCCAGCAGCTACTGAGA 597

```

Search completed: May 11, 2004, 21:52:43
Job time : 13946.8 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 10, 2004, 15:48:01 ; Search time 4264.35 Seconds

(without alignments)
3793.447 Million cell updates/sec

Title: US-09-977-053-4

Perfect score: 19973

Sequence: 1 MPRFLAFCCMGLALVSGMAT.....CHCLSSWTGHCNSKRTGCP 3571

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2941586 seqs, 2264995651 residues

Total number of hits satisfying chosen parameters: 5883172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09977053/runat_06052004_075944_18296/app_query.fasta_1.5710
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNIT8=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=25 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09977053@cgn_1_1_484@runat_06052004_075944_18296
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
14:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16:	/cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17:	/cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
18:	/cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
19:	/cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	19958	99.9	10878	9	US-09-911-842-1	Sequence 1, Appli
2	19958	99.9	10878	14	US-10-150-821-1	Sequence 1, Appli
3	19769	99.0	10674	16	US-10-295-027-429	Sequence 429, App
4	19769	99.0	10674	16	US-10-295-027-429	Sequence 1080, App
5	19616	98.2	11152	16	US-10-028-248A-7	Sequence 7, Appli
6	19616	98.2	11152	16	US-10-107-782-7	Sequence 5, Appli
7	19566	98.0	11158	16	US-10-028-248A-5	Sequence 5, Appli
8	19566	98.0	11158	16	US-10-107-782-5	Sequence 3, Appli
9	16707	83.6	11230	9	US-09-911-842-3	Sequence 3, Appli
10	16707	83.6	11230	14	US-10-150-821-3	Sequence 21, Appli
11	8427	42.2	6153	9	US-09-898-570-21	Sequence 21, Appli
12	8427	42.2	6153	10	US-09-839-446-21	Sequence 21, Appli
13	7471	37.4	3991	17	US-10-311-623-24	Sequence 24, Appli
14	6452	32.3	3804	9	US-09-764-853-214	Sequence 214, App
15	6452	32.3	3804	9	US-09-764-898-60	Sequence 60, Appli
16	6452	32.3	3804	10	US-09-764-881-21	Sequence 21, Appli
17	6452	32.3	3804	11	US-09-764-875-275	Sequence 275, App
18	6452	32.3	3804	13	US-09-764-881-21	Sequence 21, Appli
19	6452	32.3	3804	13	US-09-764-893-43	Sequence 43, Appli
20	6452	32.3	3804	15	US-10-073-865-43	Sequence 43, Appli
21	6452	32.3	3804	16	US-10-242-747-21	Sequence 21, Appli
22	5848	29.3	7286	11	US-09-984-429-525	Sequence 525, App
23	5485	27.5	3448	13	US-10-302-172-34	Sequence 34, Appli
24	4930	24.7	3262	13	US-10-466-164-29	Sequence 29, Appli
25	4058	20.3	2437	9	US-09-728-952-93	Sequence 93, Appli
26	4045	20.3	2444	9	US-09-728-952-46	Sequence 46, Appli
27	3596	18.0	1952	9	US-09-898-570-19	Sequence 19, Appli
28	3596	18.0	1952	10	US-09-839-446-19	Sequence 19, Appli
29	3133	15.7	2064	10	US-09-822-846-55	Sequence 55, Appli
30	2758	13.8	1493	9	US-09-764-853-396	Sequence 396, App
31	2758	13.8	1493	9	US-09-764-898-132	Sequence 132, App
32	2758	13.8	1493	10	US-09-764-881-79	Sequence 79, Appli
33	2758	13.8	1493	11	US-09-764-875-556	Sequence 556, App
34	2758	13.8	1493	13	US-09-764-881-79	Sequence 79, Appli
35	2758	13.8	1493	13	US-09-764-893-66	Sequence 66, Appli
36	2758	13.8	1493	15	US-10-073-865-66	Sequence 66, Appli
37	2758	13.8	1493	16	US-10-242-747-79	Sequence 79, Appli
38	2638	13.2	1652	15	US-10-146-473-17	Sequence 17, Appli
39	2305	11.5	1709	9	US-09-898-570-17	Sequence 17, Appli
40	2305	11.5	1709	10	US-09-839-446-17	Sequence 17, Appli
41	2032	10.2	7313	15	US-10-037-270-259	Sequence 259, App
42	2032	10.2	7313	16	US-10-117-722-259	Sequence 259, App
43	2027	10.1	6951	17	US-10-641-643-1265	Sequence 1265, Ap
44	2006	10.0	7028	13	US-10-376-774-802	Sequence 802, App
45	1977	9.9	1408	10	US-09-822-846-54	Sequence 54, Appli
46	1822	9.1	10136	13	US-10-016-248-1	Sequence 1, Appli
47	1606.5	8.0	5420	9	US-09-954-456-1172	Sequence 1172, Ap
48	1498	7.5	1892	9	US-09-764-898-49	Sequence 49, Appli
49	1278	6.4	680	9	US-09-764-853-173	Sequence 173, App
50	1278	6.4	680	9	US-09-764-898-122	Sequence 122, App
51	1278	6.4	680	10	US-09-764-881-14	Sequence 14, Appli
52	1278	6.4	680	13	US-09-764-881-14	Sequence 14, Appli
53	1278	6.4	680	13	US-09-764-893-37	Sequence 37, Appli
54	1278	6.4	680	15	US-10-073-865-37	Sequence 37, Appli
55	1278	6.4	680	16	US-10-242-747-14	Sequence 14, Appli
56	1184	5.9	3934	17	US-10-641-643-1066	Sequence 1066, Ap
57	1182	5.9	8010	13	US-10-016-248-3	Sequence 3, Appli
58	1009	5.1	7693	13	US-10-072-012-133	Sequence 133, App
59	1004	5.0	8091	15	US-10-101-510-86	Sequence 86, Appli
60	1004	5.0	8091	15	US-10-356-625-1	Sequence 1, Appli
61	1004	5.0	8091	16	US-10-159-563-322	Sequence 322, App
62	1001	5.0	7596	15	US-10-004-113-57	Sequence 57, Appli
63	1001	5.0	8064	15	US-10-004-113-56	Sequence 56, Appli
64	1001	5.0	8221	12	US-10-152-319A-2167	Sequence 2167, Ap
65	991	5.0	7693	13	US-10-160-497-11	Sequence 11, Appli
66	991	5.0	7693	13	US-10-348-750-11	Sequence 11, Appli
67	990	5.0	7673	9	US-09-815-925-1	Sequence 1, Appli
68	990	5.0	7673	16	US-10-294-006-1	Sequence 1, Appli
69	986.5	4.9	7615	15	US-10-004-113-59	Sequence 59, Appli
70	984.5	4.9	6728	16	US-10-190-115-3	Sequence 3, Appli
71	984.5	4.9	6728	16	US-10-369-072-3	Sequence 3, Appli
72	983	4.9	7332	9	US-09-944-849-1	Sequence 1, Appli

QY 421 LeuProAsnGlyLeuTrpSerGlySerGlySerTyrCysArgValaIgrThrCysProHis 440
DB 1271 CTACCAATGGTTGTGGTTCGGGTTACAGAGCTACTCGAGGTAAAGATGTCCTCAT 1330
QY 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1331 CTCGGCAGCGGAAACATAGCCACATCAGCTGTTCTACAAGGGAAATGTTATATAGACA 1390
QY 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1391 ACATGTTGGTTCCTGCTGATGAAGGGTACAGACTAGAGGCGAGTGATAGCTTACTTGT 1450
QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1451 CAAGGAAACAGCCAGTGGGATGGCCAGAACCCGGTGTGTGGAGGCCACTGTTCCACC 1510
QY 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
DB 1511 TTTTCAGATGCCCAAGATATCATCATATCCCTCCCACTGTGGCAAGCAGCCAGCCAAA 1570
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1571 TTTGGGACGATCTGCTATGTAAGTTCGGCCAAAGGGTTCATTTTATCTGGAGTCAAGAA 1630
QY 541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
DB 1631 ATGCTGAGATGTACCACTTCTGGAAATGATGTCGAGTTCAGGCAGCTGTGTGTA 1690
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
DB 1691 GACGTGGAGGCTCTCAAAATCAACTGTCTTAAGCAGATAGAGGCTAAGACTCTGGAACAG 1750
QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 1751 CAAGATTCTGCCAATGTTTACTCGCAGATTCCACACAGCTAAGACAACTCTGTGTA 1810
QY 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
DB 1811 GTGTGAGTCCAGCTGTTCATCCAGCTTTTACCCCACTTACCTTTTCCAGTTGGAGT 1870
QY 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
DB 1871 GCTATGCTATACACGGCAACTGACCTATCCGCAACCCAGCCAGCTGCAATTTTCCATATC 1930
QY 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
DB 1931 AAGGTTTATTGATGCAGAAACCCTGTCATAGACTGGTGACAGATCTCCACCTCCGTCG 1990
QY 661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
DB 1991 GTCTCGGAAGAGGTACATGTCGACGCTGGATGGATGAGCTCATGTTCTCAGACAACTCAGG 2050
QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyThr 700
DB 2051 GCTGAATTGGTTCATTACAGAACTATACAGAGTCAATACAGAGAGACCTTTTTCCTCAAGGGGAGACT 2110
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
DB 2111 ATAGTACAGTATACAGCCACTGACCCCTCAGGCCAATAACAGGACATGTGATATCCATATT 2170
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
DB 2171 GTCATAAAGGTTCTCCCTGTGAAATTCATTTACACTGTGAATGGGATTTATATGC 2230
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
DB 2231 ACTCCAGATTAATCTGAGTCAACTGATACATTAACTTGTGAGGGGCTATGATTTCCACA 2290
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
DB 2291 GAAAGGCTCTACGAAAGTATTATTGTGCTTATGAGATGGCGTCTGGAAACCAACATAT 2350

QY 781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
DB 2351 ACCACTGAATGGCCAGACTGTGCCAAAAGCTTTTGTCTAACACACGGGTTCAAGTCTCTT 2410
QY 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
DB 2411 GAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAGTTTTTCTGAA 2470
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
DB 2471 GCATTTGAGACGACCTTGGGAAAATGGTCCCATCAITTTGTAGTGATGCGAGAGCAT 2530
QY 841 AspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
DB 2531 GACTGCAGACTGGAGGAGAACCTGACCAAAAATATTGCTCTAGAAATATAATTATGACTAT 2590
QY 861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
DB 2591 GAAATGGCTTGTGCATTTGGACCAAGTGGCTGGGGTGCAGCTAATAGCTTGAATCTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
DB 2651 TAGCATGACTTCTCTGGACACTGTGCAAGAAACAGCCACAGCATCGGCATGCCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
DB 2711 TCACGATTAAGAAGTGGCCCAATATCTGACTATAAAATTAAGTTAAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
DB 2771 ACAGCTAGTGTGCATTTACCCGATGAAGAAATGATACCTTGAATGGGAAAATCAGCA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
DB 2831 CGACTCTCTCAGACATTTGGAACTATCACAAATAAATGAAAAGAGACTCTCAACAAAGAC 2890
QY 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
DB 2891 CCCATGTATTCTTTCAGCTTGCAATCAGAAATACTTATAGCCGACAGCAATTCATTAGAA 2950
QY 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
DB 2951 ACAAAAAGGCTTCCCTTCTGCAGACCAAGGCTCAGTGTGTGAGAGGGCGTATGTGTGTC 3010
QY 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
DB 3011 AATTCCTCTTGGAACTATTAATATCTGGAATTTTCACTGTGAAGCTGTCCGGATC 3070
QY 1021 GlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
DB 3071 GGATCTCTATCAAGATGAAGAGGGCAACTTGAGTGCAGCTTTTGGCCCTCTGGGATGTAC 3130
QY 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
DB 3131 ACAGATATATATCCATTCAGAAACATCTCTGAATGTAAAGCTCAGTGTAAACAAGGCACC 3190
QY 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
DB 3191 TACTCATACAGTGGACTTTGAGACTTGTGATTCGTCCACTGGGCACTTATCAGCAAAA 3250
QY 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
DB 3251 TTTGGTTCCCGAGCTGCTCTCTGTGTCCAGAAAACACCTCAACTGTGAAAAGAGGAGCC 3310
QY 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
DB 3311 GTGAACATTTCTGCATGTGGAGTTCTTGTGTCCAGAGGAAAATTTCTCCGTTCTCGGTTA 3370
QY 1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
DB 3371 ATGCCCCGTACCCCATGCTCTCGTACATATCCACCTAATGACGAGGAGGCCCTTCTGC 3430
QY 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160

Db 3431 CTGGCTGTCCTTTATATGGAATCACTACCCATTCGCTGTTCCAGATTCATACAGATGT 3490
Qy SerSerPheSerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu 1180
Db 3491 TCAAGTTTATGTTCAACTTTCTCAGCGGCAGAGGAAGTGTGTCGCCCTCTCTT 3550
Qy GlyHisIleLeuLeuArgHisGluIleSerSerGluValPheHisGluCysPhePheAn 1200
Db 3551 GGACATATATAAAGAGGATGAATATCAGCAGTCAAGTTTTCATGAATGCTTCTTAAC 3610
Qy ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyThrValCysLeuCysPro 1220
Db 3611 CCTTCCCAATATGATGGAACTGCCAGCACTGGCGGTGTATGTTGTCTCTGTCCA 3670
Qy LeuGlyThrGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1240
Db 3671 CTTCGATATACAGGCTTAAAGTGTGAACAGATCGATGAGTGCAGGCCACTGCTTGC 3730
Qy LeuAsnAsnGlyValCysLeuAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
Db 3731 CTCAACATATGAGTTTGAAGACCTAGTTGGGAAATTCATTTGTGAGTGCCTCAGGT 3790
Qy TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsn 1280
Db 3791 TACACAGGTACCGGTGTGAAGAAATATAATAGATGAGTGTCCAGTCTTGTGTTAAAT 3850
Qy LeuGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLeuGlyPheVal 1300
Db 3851 AAAGCAATCTGTTGATGTGTGCTGCTGCTATCGTTGCACATGTGTGAAGGATTTGTA 3910
Qy GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAla 1320
Db 3911 GGCCTGCATTTGTGAACAGAACTAAATGAATGCCAGTCAAACTCATGCTTAAATATGCA 3970
Qy ValCysGluAspGlnValGlyGlyPheLeuCysLeuCysValProGlyPheLeuGlyThr 1340
Db 3971 GTCTGTGAAGACAGGTGGGGATTTCTGTGCANAATGCCACCTGGATTTTGGGTACC 4030
Qy ArgCysGlyLeuAsnValAspGluCysLeuSerGlnProCysLeuAsnGlyAlaThrCys 1360
Db 4031 CGATGTGAAGAAACCTGCATGAGTGTCTCAGTCAGCCATGCAAAAATGGAGTCACTGT 4090
Qy LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrClySerHisCys 1380
Db 4091 AAAGACGGTCCAAATAGCTTCAGATGCCCTGTGTGCAGCTGGCTTCACAGGATCACACTGT 4150
Qy GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db 4151 GAATTTGAACATCAATGAATGTCACTCTAATCCATGTAGAAATCAGGCCACCTGTGGAT 4210
Qy GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db 4211 GAATTTAATTCATACAGTTTGAATGTCAAGCCAGGATTTTCAGGCAAAAGGTGTGAACA 4270
Qy GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrClyThrValMet 1440
Db 4271 GAACAGTCTACAGGCTTTAACTCGAATTTTGAAGTTTCTGGCATCTATGGATATGTCAAG 4330
Qy LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer 1460
Db 4331 CTAGATGGCATGCTCCCATCTCTCCATGCTCTTAACCTGTACCTTCTGGATGAATCCTCT 4390
Qy AspAspMetAsnTyrGlyThrProLysTyrAlaValAspAsnGlySerAspAsnThr 1480
Db 4391 GACGACATGAACATATGAACACCAATCTCTATGCAATGATTAAGCGCAGCGCAATATCC 4450
Qy LeuLeuLeuThrAspTyrAsnGlyTrpValLeuTyrValAsnGlyArgGluLeuValThr 1500
Db 4451 TTGCTCCTGACTGATATTAACGGCTGGGTCTTTATGTGAATGGCAGGGAAGATTAACA 4510
Qy AsnCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAla 1520

Db 4511 AACTGTCTCCTCGGTGAATGATGGCAGATGGCATATATTGCAATCACTTGGCAAGTGCC 4570
Qy AsnGlyIleIleTrpLeuValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db 4571 AATGGCATCTGGAAAGTCTATATCGATGGGAAATATATCTGAACGGTGTGCTGGCCTCTCT 4630
Qy ValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLys 1560
Db 4631 GTTGGTTTGGCCATACCTGCTGGTGTGGTTAGTTCTGGGGCNAGAGCAAGCAAAAAA 4690
Qy GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp 1580
Db 4691 GGAGAGGATTCAGCCAGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG 4750
Qy AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db 4751 GACTATGTCTGTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTCCGCCAGAGGAATC 4810
Qy SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db 4811 AGTAAAGGAACCTGTTAGCATGGCTGATTTCTTGTCAAGGAATGTGGGAAAGTGAAG 4870
Qy IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 4871 ATCGATTCTAAGCAGCATATTTTGTCTGATTGCCACGCTTAGGAGGGTCAGTGCCTCAT 4930
Qy LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 4931 CTGAGAACCTGCATCTGAAGATTTAAAGCCAGGTTCCAAGTCAATCTGTTGTGTATCCA 4990
Qy GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln 1680
Db 4991 GGTTCAGCTGCTGGGAACCTGTGTGAGTACTGTCTGAATCAAGACAGTGCACACAA 5050
Qy ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db 5051 CCACTTCCTCACTGTGAACGCAATAGTCTGGGTGCCACCTCTCTTTGGAGATGGCTTC 5110
Qy HisSerAlaAspAspPheTyrAlaGlySerThrValTyrGlnCysAsnAsnGlyTyr 1720
Db 5111 CATTCAAGCGATGACTTCTATCTGGCAGCAGTAACCTTACCAGTGCACCAATGGCTAC 5170
Qy TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer 1740
Db 5171 TATCTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGTTTCA 5230
Qy ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5231 CCATCTGCTCTGATGTGATGAGTGTGAGTGTGATCAGATTTGTAGTGAGCATGCTCT 5290
Qy CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db 5291 TGCCTGAACGTAGATGGATCTTACATATGTTTCAATGTGTCCCACCGTACACAGAGATGGG 5350
Qy LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5351 AAAAACTGTGCAGAACCTATAAATGTAAGGCTCCAGGAATCCGGAATGSCCACTCC 5410
Qy SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db 5411 TCAGGTGAGATTTATACAGTAGTGTGCCAAGTCACATTTTCTGTGTGAGGAAGATACCAG 5470
Qy LeuMetGlyValThrIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 5471 TTGATGGGAGTAAACAAATCAATGTTTGGAGTCTGGAGATGGATTCATCTANTACCA 5530
Qy TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
Db 5531 TATTTGAAGCTGTTTCAATGTGTAAACCGGCTATTCCAGAAAAATGCTTGCATTGAGGAG 5590
Qy LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla 1880
Db 5591 TTAGCATTTACTTTTGGCAGCAAGTGAATATAGGTGTAAATAAGGATATATCTCTGGCC 5650

Db 7811 TGGTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCTCCCTCTCATATAGAT 7870
Qy PheGlyAspCysThrIysLeuIysAspGlnGlyTyrPheGluGlnGluAspMet 2640
Db 7871 TTGGAGAGCTGTACTAAACTCAAGATGACACAGGAGATATTTTGGACAGAGACGACATG 7930
Qy MetGluValProTyrValThrProHisProProTyrHisLeuGlyValaValalalyenThr 2660
Db 7931 ATGGAGTTCATATGTGACTCTCCACCTCTTATCATTTGGAGAGTGTCTAAACCC 7990
Qy TrrGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhelLeuTyrGlyThrMet 2680
Db 7991 TGGGAAATACAAAGAGTCTCTGCTACACATTCAAACTTCTGTATGGTACCATG 8050
Qy ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuileCysGln 2700
Db 8051 GTTTCATACACCTGTAATCCAGGATGAATCTCTGGGGAACCTCTGTCTGATCTGCCAG 8110
Qy GluAspGlyThrTrrAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
Db 8111 GAAGATGGAACCTTGAATGGCAGTGCACCATCTCTGCATTTCAATTGAATGTGACTTGCT 8170
Qy ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln 2740
Db 8171 ACTGCTCTGAAATATGGCTTTTGTGCTTTTACAGAGACTAGCATGGGAGTCTGTGAG 8230
Qy TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
Db 8231 TATAGCTGTAAACCTGGACACATCTAGTGGGCTCTGACTTAAGGCTTGTGTAGAGAT 8290
Qy ArgIysThrProSerGlyValSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
Db 8291 AGAAAGTGAGTGGTGGCTCCCGACGCTGTGAAGCCATTTTCAGCAAAAGCAAAATCCA 8350
Qy ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu 2800
Db 8351 GTCATGAATGGATCCATCAAGAGAGCACTACACATCTGACGACGTTGTACTATGAG 8410
Qy CysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn 2820
Db 8411 TGTGACCCCGGATATGTGTGAATGGCACTGAGAGAGAGACATGCGCAGGATGACAAAAAC 8470
Qy TrrAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla 2840
Db 8471 TGGGATGAGGATGAGCCCATTTGCATTCCTGTGGACTGCGATTCACCCCGCTCAGCC 8530
Qy AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn 2860
Db 8531 AATGGCCAGGTGAGAGAGACGAGTACACATTTCCAAAGAGATTCATACACTTGCAT 8590
Qy GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrrSer 2880
Db 8591 GAAGGGTCTTGCTTGGAGGAGCCAGGAGTGGGTGTGTCTTGTCCCAATGGAAGTTGGAGT 8650
Qy GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly 2900
Db 8651 GGAGCCATCCCGACTGTGTGCTGTGCATGTGCCACCCCGCCCAACTGGCCAAATGGG 8710
Qy ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly 2920
Db 8711 GTGACGGAGGCTTGACTATGGCTTCATGAAGGAGTAAACATTCACCTGTCTATGAGGC 8770
Qy TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrrAspAlaGlu 2940
Db 8771 TACATCTTGACGGTGTCTCAAAACTCACTGTCTGCTGAGTGCAGATGGCACTGGGATGCAG 8830
Qy IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
Db 8831 ATTCTCTCTGTAAACCACTCAACTGTGGACCTCTCTGAAGATCTTGGCCCATGGTTCCCT 8890
Qy AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980

Db 8891 AATGGTTTTTCTTTATTTCATGGGGCCATATACAGTATACAGTCTTCTCTGGTTATAAG 8950
Qy LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrrSerGlySerSerPro 3000
Db 8951 CTCCATGGAAATTTCAATCAAGAAAGTGCCTCTCCATATGGTCTCTGGAGTGGCAGCTCACCT 9010
Qy SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr 3020
Db 9011 TCTTCCCTGCCTTGCATGTTCACACACAGTAATTAATATGAATATGAATGTCTCATATGGACA 9070
Qy AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
Db 9071 GATTTTGAAGTGTGAAGAGGAGCCCGGATTCAGTGTCTCAAGGCTTCAAGCTCTTAGGA 9130
Qy LeuSerGluIleThrCysGluAlaAspGlyGlnTrrSerSerCysGlyPheProHisCysGlu 3060
Db 9131 CTTTCTGGAATCACCTGTGAAGCCGATGGCAGTGGAGTCTGGGTTCGCCCACTGTGAA 9190
Qy HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer 3080
Db 9191 CACACTTCTTGTGGTCTCTTCCATATGATACCAATGCCGTTTCATCAGTGGACCACTCT 9250
Qy TrrLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
Db 9251 TGGAAAGGAAAATGTGATAACTTACAGCTGCAGTCTGGATATGTCTATACAGGCGAGTTCA 9310
Qy AspLeuIleCysThrGluLysGlyValTrrSerGlnProTyrProValCysGluProLeu 3120
Db 9311 GATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCAGTCTGTGTAGCCCTTG 9370
Qy SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr 3140
Db 9371 TCCTGTGGGTCCCGACCGTCTGTCCCAATGTCAGTGGCAACTGGAGAGGCACACACTAT 9430
Qy GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThr 3160
Db 9431 GAAAGTGAAGTGAATACTCAGATGTCTGGAAGGTATACGATGATACAGATACAGATACA 9490
Qy PheThrCysGlnLysAspGlyArgTrrPheProGluArgIleSerCysSerProLysLys 3180
Db 9491 TTCACCTGTGAGAAAGTGTGCTGGTGTCTCTGAGAGAAATCTCTGCGAGTCTCTAAAAA 9550
Qy CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
Db 9551 TGTCTCTCTCCCGAAAACATACATATATCTTGTTCATGGGAGCGATTTCACTGTGAT 9610
Qy ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
Db 9611 AGGCAAGTTTCTGTCTCATGTGCAGAAAGGTATACCTTTGAGGGAGTTAAACATATCAGTA 9670
Qy CysGlnLeuAspGlyThrTrrGluProProPheSerAspGluSerCysSerProValSer 3240
Db 9671 TGTCACTGTATGGAACTTGGAGCCACCATTTCTCCGATGAATCTTTGCGAGTCCAGTTCT 9730
Qy CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
Db 9731 TGTGGGAAACCTGAAGTCCAGAACATGATGATTTGTGTTGGCAGTAAATACACTTTGAA 9790
Qy SerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgVal 3280
Db 9791 AGCACAAATATTATTCAGTGTGAGCTGCTGATGAATAGAGGGGAAACAGGGAAACCGTC 9850
Qy CysGlnGluAsnArgGlnTrrSerGlyValAlaIleCysLysGluThrArgCysGlu 3300
Db 9851 TGGCAGAGGAAACAGACAGTGGAGTGGGGGTGGCAATATGCAAGAGACAGGTTGAA 9910
Qy ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsn 3320
Db 9911 ACTCCACTTGATTTCTCAATGGGAAAGCTGACATTTGAAACAGGACGACTGGACCCAAC 9970
Qy ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
Db 9971 GTGGTATATTCCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCGACACTGACA 10030

Qy 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyLysSerSerCys 340
 Db 971 ACAGCTGGCCCATCGGGACATACAACTCAAGGCTCACAGGAGGAATCAGCAGTTGC 1030
 Qy 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
 Db 1031 ATTCATGTCTCTGATGAATAATCAACCTCTCACCTGGAGGACATCCCTCGAAGACTGT 1090
 Qy 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
 Db 1091 GTCTGCAGAGAGGATACAGGGCATCTGGCCAGACTGTGAATGTGTCCATGCGCTGCC 1150
 Qy 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
 Db 1151 CTGAAGCCTCCCGAAATGGTTACTTTATCCAAACACTTGCACCAACCACTTCAATGCA 1210
 Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
 Db 1211 GCCTGTGGGGTCCGATGTCACTGGATTTGATCTGTGGGAAGCAGCATCATCTTATGT 1270
 Qy 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
 Db 1271 CTACCAATGGTTGTGGTCCGGTTTCAGAGAGCTACTCGAGAGTAAGAACAATGCTCTCAT 1330
 Qy 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
 Db 1331 CTCGCCACGCCGAACATGGCCACATCAGCTGTTCTACAAAGGGAATGTTATTAAGACA 1390
 Qy 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
 Db 1391 ACATGTTTGGTTGGCTGTGATGAAGGGTACACACTAGAAGGCAGTGATGAAGCTTACTTGT 1450
 Qy 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
 Db 1451 CAAGNAACAGCCAGTGGATGGGCGAGNACCCCGGTGTGTGGAGCGGCACATGTTCCACC 1510
 Qy 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProLys 520
 Db 1511 TTTTCAGATGCCCAAGATATCATCATATCCGCCACACTGTGGCAGCAGCCAGCCAAA 1570
 Qy 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
 Db 1571 TTTGGAGCATCTGCTATGATTAAGTTGCGCCCAAGGGTTCAATTTATCTGGAGTCAAGAA 1630
 Qy 541 MetLeuArgCysThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLys 560
 Db 1631 ATGCTGAGATGTACCCTTCTGGAAATATGGAATGTGGAGTTTCAGGCAGCTGTGTGTA 1690
 Qy 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
 Db 1691 GACGTGGAGGCTCTCTCAATCAACTGTCTTAAGGACATAGAGGCTTAAGACTCTGGAACAG 1750
 Qy 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
 Db 1751 CAAGATTCTGCCAATGTTTACCTGGCAGATTCACACAGCTTAAGACAACTCTCTGGTGA 1810
 Qy 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
 Db 1811 GTGTGAGTCCAGTTCATCCAGCTTTCACCCCACTTACTCTTCCAGTGGAGATGTT 1870
 Qy 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
 Db 1871 GCTATCGTATACAGGCAACTGACCTATCCGGCAACACAGGCCAGCTGCTATTTCCATATC 1930
 Qy 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
 Db 1931 AAGGTTATGTATGCAGAACCCCTGTCTAGACTGTGTGAGATCTCCACTCCCGTCCAG 1990
 Qy 661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
 Db 1991 GTCTCGGAGAGGTATATGTCGGCAGCTGGGATGAGCCTCAGTTCTTCAGACAACTCAGG 2050

Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
 Db 2051 GCTGAATTTGGTCATTTACCAAGATCATACCAAGGAGACCTTTTCCCTCAAGGGAGACT 2110
 Qy 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
 Db 2111 ATAGTACAGTATACGCCACTGACCCCTCAGGCATTAACAGGACATGTGATATCCATAT 2170
 Qy 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
 Db 2171 GTCATAAAGGTTCTCCCTGTGAATTCATTCACACCTGTAAATGGGAGTTTATATGC 2230
 Qy 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
 Db 2231 ACTCCAGATTAATACCTGGAGTCACTGTACATTAATCTGCTGGAGGCTATGATTTTACA 2290
 Qy 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
 Db 2291 GAAGGCTCTACTGACAGATTAATTTGTGCTTATGAAGTGGCGTCTGGAAACCAACATAT 2350
 Qy 781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
 Db 2351 ACCACTGAATGGCCAGACTGTGCCAAAAAAGCTTTTGTCTAACCCAGGTTCAAGTCTCTT 2410
 Qy 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
 Db 2411 CAGATGTTCTACAAAGCAGCTGTTGTGATGACACAGATCTGATGAAGAGTTTTCGTAA 2470
 Qy 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
 Db 2471 GCATTTTGAGACACCTCGGAAAAAATGTGCCCATCATTTTGTAGTATGACAGGACATT 2530
 Qy 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
 Db 2531 GACTGCAGACTGGAGAGAACCTGCACAAAAAATATTGCTTAGATATATTAATTAAGTAT 2590
 Qy 861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValaAlaAsnArgLeuAspTyrSer 880
 Db 2591 GAAATATGGCTTTGCAATTTGGACAGGTGCTGGGGTGCAGCTAATAGGCTGGATTACTCT 2650
 Qy 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
 Db 2651 TAGCATGACTCTCTGGACACTGTGCAAGAAACAGCCACCAAGCATCGCAATGCCAAGTCC 2710
 Qy 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
 Db 2711 TCACGGATTAAAGAAAGTCCCCCATTTACTGACTATAAAAAATTAAGTTAAATTTAAACATC 2770
 Qy 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
 Db 2771 ACAGCTAGTGTGCTCATTTACCCGATGAAGAAATGATACCCCTTGAAATGGGAAATTCAGCAA 2830
 Qy 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
 Db 2831 CGACTCTTCAGACATTTGGAACACTATACAAATAAATGAAAGGACTCTCAACAAAGAC 2890
 Qy 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
 Db 2891 CCCATGTAATCTCTTCAGCTTGCAATCAGAAATATCTTATATAGCCGACAGCAATTCATTAGNA 2950
 Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
 Db 2951 ACAAAAAGGCTTCCCTCTTCGACACCAAGGCTCAGTGTCTGAGAGGCGGTATGTGTGTC 3010
 Qy 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
 Db 3011 AATTGCCCCCTTTGGAAACCTTATTATATCTGGAACATTTTACCTGTGAAAGCTGCCGATC 3070
 Qy 1021 GlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
 Db 3071 GGATCTCATCAGATGAAGAGGCAACTTGTAGTGCAGCTTTTGGCCCTCTCGGAGATGTAC 3130
 Qy 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060

Db 3131 ACAGATATATATCCCTTTACAGAAACAATCTCGATTGTAAAGCTCAGTGTAACACAGGACCC 3190
Qy TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
Db TACTCATACAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCACCTTATCAGGCCAAAA 3250
Qy PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
Db TTTGGTTCGCGAGCTGCCTCTCGTGTCCAGAAACACCTCAACTGTGAAAGAGGAGCC 3310
Qy ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
Db GTGAACATTTCTGCATGTGGAGTCTCTGTTCAGAGGAAATTTCTGCGTTCGGSYTA 3370
Qy MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
Db ATGCCCTGTCAACCATGTCTCGTGAATATACCACTTAATCAGGAGGAGGCTTCTGCG 3430
Qy LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
Db CTGCGCTGTCCCTTTATGGAACTTACCCCAATTCGCTGGTTCCAGATCCATCAGCAATGT 3490
Qy SerSerPheSerThrPheSerAlaAlaGluSerValProProAlaSerLeu 1180
Db TCAAGTTTATAGTTCAACTTCTCAGCGGCGAGGAAAGTGTGGTCCCTCTGCTCTCT 3550
Qy GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
Db GGACATATTAAAAAGAGGCATGAATCAGCAGTCAGGTTTTTCCATGAATGCTTCTTTAAC 3610
Qy ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
Db CCTTGGCCACAATAGTGGAACTCTCGCAGCAACTTGGGCGTGGTTATGTTTGTCTCTGCCA 3670
Qy LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
Db CTGGATATACAGCTTAAAGTGTGAACACAGACATCGATGAGTGCACCCACTGCTTGC 3730
Qy LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
Db CTCACAAATGGAGTTTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGCCCATCAGGT 3790
Qy TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
Db TACACAGGTCAGCGGTGTGAAGAAATATATAATGAGTGTAGTGCAGCTCTTGTGTTAAAT 3850
Qy LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
Db AAAGGAATCTGTGTTGATGTTGGTGGCTGCTATCGTGCACATGTTGTGAAGGATTTGTA 3910
Qy GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAla 1320
Db GGCTCGCATTTGTCAAAACAGAAATCAATGAATGCGAGTCAAAACCCATCTCTTAAATAATGCA 3970
Qy ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
Db GTCTGTGAAGACAGGTTGGGGATTTCTGTGCAAAATGCCACCTGATTTTGGGTPACC 4030
Qy ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
Db CGATGTGGAAGAACCTCGATGATGATGCTCAGTCAGCCATGCCAAAATGAGGCTACCTGT 4090
Qy LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db AAAGACGGTGCCAAATAGCTTCAGATGCTGTGTGCAGCTGGCTTCAAGGATCACACTGT 4150
Qy GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db GAATTGAACATCAATGAATGCTAGTCTTAATCCATGTAGAAATCAGGCCACCTGTGTGGAT 4210
Qy GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420

Db 4211 GAATTAATAATTCATACAGTGTGTAAATGTCCAGCAGGATTTTTCAGGCCAAAAGGTGTGAACA 4270
Qy GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db GAACAGTCTACAGGCTTTAACTCGATTTTGAAGTTTCTGGCATCTATGGGATATGTCTATG 4330
Qy LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSer 1460
Db CTAGATGGCATGTCTCCCATCTCTCCATGCTTAACCTGTACTCTTCTGGATGAAATCTCTCT 4390
Qy AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db GAGCAGCATGAATATGGAACACCAATCTCTATGCAGTTGATAACCGCAGCGACAAATACC 4450
Qy LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db TTTGCTCTGACTGATATATAACGGCTGGGTCTTTATGTGAATGGCAGGGAAGAATAACA 4510
Qy AsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSerAla 1520
Db AACTGTCCCTCGGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGCAAGTGGC 4570
Qy AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db AATGGCATCTGGAAGTCTATATCGATGGGAAATTTATCTGACGGTGTGTGCTGGCCTCTCT 4630
Qy ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysLys 1560
Db GTTGGTGTGGCCATACCTGGTGGTGGTGTGCTGTCTGGGGCAAGACAGCAAGAAAAA 4690
Qy GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
Db GGAGAGGAGATTCAGCCACAGCTGAGTCTTTGTGGGCTCCATAGCCAGCTCAACTCTGG 4750
Qy AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db GACTATGTCTGTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTCTGCCCAGAGGAATC 4810
Qy SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db AGTAAAGGAACCGTTAGCATGGCCCTGATTTCTTGTTCAGGAATTTGTGGGGAAGTGAAG 4870
Qy IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db ATCGATTTCTAAGACCATATTTTGTCTGATTTGCCCACTAGGAGGTCAGTGGCTCAT 4930
Qy LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db CTGAGAACTGCATCTGAAGATTTAAAGCCAGGTTTCAAAAGTCAATCTCTTCTGTGATCCA 4990
Qy GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
Db GGCTTCAGCTGGTGGGAAACCCCTGTGCAGTACTGTCTGAATCAAGGACAGTGGACACA 5050
Qy ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db CCACCTCTCTCTGTAAGCATTTAGCTGTGGGTGCCACCTCTCTTTCAGGAATGGCTTC 5110
Qy HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db CAITCAGCCGATGACTTCTATGTCTGGCAGCAGTAACTTACCAGTGCACAAATGGCTAC 5170
Qy TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSer 1740
Db TATCTATTGGTGGTGACTCAGGATGTTCTGTACAGATATGGGAGCTGGACGGCTGTCA 5230
Qy ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db CCATCTCTCCCTTGTATGTGATGAGTGTGAGTTCGATTCGATTCAGATTTGTAGTGCATGCTTCT 5290
Qy CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db TGCCTGAACCGTAGATGGATCCTACATATATGTTTCTATGTGTCCACCGCTACACAGGAGATGG 5350

QY 1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
DB 5351 AAAAACTGTGCAGAACCTATAAAATGTAGGCTCCAGGAAATCCGAAAAATGCCACTCC 5410
QY 1801 SerGlyGluIleThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
DB 5411 TCAGGTGAGATTATACAGTAGTGGCCGAGTCACATTTTGTGTGTCAGGAGGATACAG 5470
QY 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluThrPheAsnHisLeuPro 1840
DB 5471 TTGATGGGAGTAACCAAAATCACATGTTGGAGTCTGGAGAAATGGAATCATCTAATACCA 5530
QY 1841 TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
DB 5531 TATTGTAAAGCTGTTTCATGTGGTAAACCGGCTATTCCAGNAAAATGTTGCAATTGAGGAG 5590
QY 1861 LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla 1880
DB 5591 TTAGCATTTACTTTTGGCAGCAAGTGACATATAGGTGTAAATAGGATATATCTGGCC 5650
QY 1881 GlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProValCys 1900
DB 5651 GGTGATAAGAAATCATCTGTCTTGCTACAGTCTCTGGAGTCATTTCCCTCTCTGTGTGT 5710
QY 1901 GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly 1920
DB 5711 GAACCAAGTGAAGTGTCTAGTCCGGAATAATAAATAATGGAAAAATATATTTTGAAGTGG 5770
QY 1921 LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro 1940
DB 5771 CTTACTACCTTTCTACTGTCATCATATTCATGCGATACAGGATACAGCTTACAGGGCCCT 5830
QY 1941 SerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeu 1960
DB 5831 TCCATTATTGAATGACGGCTTCTGGCATCTGGGACAGAGCGCCACCTGCTGTGCACCTC 5890
QY 1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
DB 5891 GTCTTCTGTGGAGAACCACTCCCATCAAAAGATGCTGTCAATACCGGGAAATTAATTCAC 5950
QY 1981 PheArgAsnThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr 2000
DB 5951 TTCAGAACACCGTCACTTACACTTGCAGAAAGAGGTATACTCTGTCTGTGTGCACCC 6010
QY 2001 IleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSer 2020
DB 6011 ATTGAATGCTTGGCCGACGGCAAGTGGAGTAGAAGTGACACAGCTGCTGTCTCTCC 6070
QY 2021 CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly 2040
DB 6071 TGTGATGAGCCACCACTTGTGGACCAACGCTCTCCAGAGACTGCCCATCGGCTCTTTGGA 6130
QY 2041 AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu 2060
DB 6131 GCAATTGCAATTTACTACTCTCTGTGTTGATAGCTTAGCGTAGACACAAATTCACAGCTTCTC 6190
QY 2061 CysAsnAlaGlnGlyLysTrpValProProGluGlyLysAspMetProArgCysIleAla 2080
DB 6191 TGCAATGCCAGGCAAGTGGGTATCCCCCAGAAAGGTCAAGACATGCCCCGTTGTATAGCT 6250
QY 2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
DB 6251 CATTTCTGTGAAAAACCTCATCGGTTCCTATAGCATCTTGGAAATCTGTGAGCAAAAGCA 6310
QY 2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
DB 6311 AAATTTGACGTGGCTCAGTGTGAGCTTTAAATGATGGAAGGCTTTGTACTGACACACC 6370
QY 2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
DB 6371 TCAGCAAGATTGAATGATGAGAGTGGGAGTGGAAACCTTCCCCCATGTCCATCCAG 6430

QY 2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
DB 6431 TGCATCCCTGTGCGGTGTGGAGAGCCCAAGCATCATGAATGGCTATGCAAGTGGATCA 6490
QY 2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
DB 6491 AACTACAGTTTTGGAGCCATGTGTGGCTTACAGCTGCAACAAGGGTTCTACATCAAGGG 6550
QY 2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis 2200
DB 6551 GAAAAAGAGAGACCTCGCAAGCCACAGGGCAGTGGAGTAGTCTTATACGAGCTGGCAC 6610
QY 2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
DB 6611 CCGTATCTTGTGTGGAACCACTTAAGGTGAGAAATGGCTTTCTGAGCATACACTGGC 6670
QY 2221 ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer 2240
DB 6671 AGGATCTTTGAGAGTGAAGTGAAGTATCAGTGTAAACCGGGCTATAAGTCAGTCGGAAGT 6730
QY 2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
DB 6731 CCTGTATTGTCTGCCAAGCAATCGCCACTGGCACAGTGAATCCCTCTGTGATGTGTT 6790
QY 2261 ProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe 2280
DB 6791 CCTCTGACTGTGAAAAACCTCCCGATCCAGATGGCTTCATGAAGAGGAAACCTTT 6850
QY 2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
DB 6851 GAAGTAGGCTCCAAGGTTCAAGTTTTCTCTAATAGGGTTATGAGCTTGTGTGTGACAGT 6910
QY 2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320
DB 6911 TCTTGACATGTCCAGAAATCTGGCAATGGAAATGAAGTCAAAATCCAAATGCGATGCT 6970
QY 2321 AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLysGluLeuThrThr 2340
DB 6971 GCCAGTGGCCAGAGCGCCCTCTTGGAAACCAAGCTAGTATTAAGAGGTTGACACC 7030
QY 2341 GluValGlyValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
DB 7031 GAGTAGGAGTTGTGACATTTTCTGTAAAGAGGGCATGTCTGCAAGGCCCTCTGTC 7090
QY 2361 LeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
DB 7091 CTGAAATGCTTGCATCCAGCAATGGAAATGACTCTTCTCCCTGTTGTGAAGATTGTTCT 7150
QY 2381 CysThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
DB 7151 TGTACCCCACTCCCTAATTTCTTTGGTGTGCCCATTTCTTCTCTGCTCTTCATTT 7210
QY 2401 GlySerThrValLysTyrSerCysValGlyPhePheLeuArgGlyAsnSerThrThr 2420
DB 7211 GGAAGTACTGTCAAGTATTTCTTGTGTAGTGGTTTTTCTTAAGAGGAAATTTACACC 7270
QY 2421 LeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys 2440
DB 7271 CTCTGCCAACCTGATGGCACCTCGAGCTCTCACTGCCAGAAATGTGTTCAGTAGAATGT 7330
QY 2441 ProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlnLeuAlaTyrLeuSer 2460
DB 7331 CCCCACCTGAGGAAATCCCAATGGATCATTTGATGTGCAAGGCCCTTGCCTATCTCAGC 7390
QY 2461 ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys 2480
DB 7391 ACAGCTCTCTATACCTGCAAGCCAGGCTTTGAATGGTGGGAAATACTTACCACCTTGT 7450
QY 2481 GlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLys 2500
DB 7451 GGAGAAAAATGGTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTGTAGTGGCTGAAA 2520
QY 2501 ProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal 2520

DB 7511 CCNAGGAGATTGTAATGGCAATTTCTTTACCGGACCTACACTATGGACAGACGGTT 7570
QY 2521 ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu 2540
DB 7571 ACCTACTCTTGCACACGAGGCTTTGGCTCGAAGTCCCAAGTCCCACTGACCTGTAGAG 7630
QY 2541 ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln 2560
DB 7631 ACAGGTGATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7690
QY 2561 ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleTyr 2580
DB 7691 CCATTTGAAATGGTTTTGTAGAGGTGCAGATTACAGCTATGCTGCTGATCAATCATCTAC 7750
QY 2581 SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly 2600
DB 7751 AGTTGCTTCCCTGGGTTTTCAGGTGGCTGGTTCATGCGCATGCGACCTGTGAAGAGTCAGGA 7810
QY 2601 TrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp 2620
DB 7811 TGCTCAAGTTCCATCCCAACATGATGCCAATAGACTGTGGCTTCCCTCTCATATAGAT 7870
QY 2621 PheGlyAspCysThrIleLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMet 2640
DB 7871 TTTGGAGACTGTACTAACTCAAGATGACAGGGATATTTTGGGCGAGAGAGACATG 7930
QY 2641 MetGluValProTyrValThrProHisProTyrHisLeuGlyAlaValAlaLysThr 2660
DB 7931 ATGGAAATGTCATATGTGACTCTCCCTCCCTTATCATTTGGGAGCAGTGGCTAAAAAC 7990
QY 2661 TrpGluAsnThrIleGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet 2680
DB 7991 TGGGAAATACAAAGAGTCTCTGCTTACATTCATCAAACTTTCTGTATGTGTACCATG 8050
QY 2681 ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGln 2700
DB 8051 GTTTCATACACCTGTATATCCAGATATGAATCTCTGGGGAACCTGTGTGATGTGCCAG 8110
QY 2701 GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
DB 8111 GAAGATGGAACTTGGAAATGGCAGTGCACCATCTGCATTTCAATTTGAATGTGACTTGCCT 8170
QY 2721 ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln 2740
DB 8171 ACTGCTCCTGAAATGCTTTTTCGTTTTTACAGAGACTAGCATGGAGTGTGTGCAG 8230
QY 2741 TyrSerCysIleProGlyHisIleLeuAlaGlySerAspLeuArgLeuGluAsn 2760
DB 8231 TATAGCTGTAACTGGACACATTTAGTGGCTCTGACTTAAGCTTTGTCTAGAGAT 8290
QY 2761 ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysIleLysProAsnPro 2780
DB 8291 AGAAGTGGAGTGGTGGCTCCCGCTGCTGGAAGCCATTTTCATGCAAAAGCCAAATCCA 8350
QY 2781 ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrGlu 2800
DB 8351 GTCATGAATGGATCCATCAAGAGAGCAGTACATACATCTGAGCAGCTGTACTATGAG 8410
QY 2801 CysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn 2820
DB 8411 TGTGACCCCGATATGTGCTGAATGGCAGCTGAGAGAGACATGCCAGGATGACAAAAC 8470
QY 2821 TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla 2840
DB 8471 TGGGATGAGGATGAGCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8530
QY 2841 AsnGlyGlnValArgGlyAspGluTyrThrPheGlnIleGluTyrThrCysAsn 2860
DB 8531 AATGGCCAGGTGAGAGAGCAGTACATTCATTCAAAAGAGATTGAATACACTTGCAT 8590
QY 2861 GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer 2880

DB 8591 GAAGGGTTCTTGCTTGAGGAGGCGCAGGAGTCGGGTTTGTCTTGCCAATGGAAGTTGGAGT 8650
QY 2881 GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly 2900
DB 8651 GGAGCCACTCCCACTGT 8710
QY 2901 ValThrGluGlyLeuAspTyrGlyPheMetLysLeuValThrPheHisCysHisGluGly 2920
DB 8711 GTACGGAAGGCTTGACTATGCTTTCATGAGGAGTAACATTCACCTGTCTCATGAGGGC 8770
QY 2921 TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu 2940
DB 8771 TACATCTTGACGGTGTCTCCAAACTCACCTGTCACTCAGATGCAATGCGGATGCGAG 8830
QY 2941 IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
DB 8831 ATTCTCTCTGTATAACCACTGCTGAGACCTCTCAAGATCTTGCCCATGGTTTCCCT 8890
QY 2961 AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980
DB 8891 AATGGTTTTCTTTATTCATGGGGCCATATACAGTATCAGTGTCTTCTGTGTATTAAG 8950
QY 2981 LeuHisGlyAsnSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerPro 3000
DB 8951 CTCATGGAAATTCATCAAGAGGTGCTCTCCAATGGCTCTGAGTGGCAGCTCACT 3010
QY 3001 SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr 3020
DB 9011 TCTGCTGCTCTTTCAGATGTTCACACCACTGAATTAATGAATGAGACTGTCAATGGGACA 3070
QY 3021 AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
DB 9071 GATTTGACTGTGGAAAGGCGACCCGGATTCAGTGTCTCAAAAGCTTCAAGCTCTCTAGSA 9130
QY 3041 LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu 3060
DB 9131 CTTTCTGAATCACCCTGTGAAGCCGATGCGCAGTGGAGCTCTGGGTTCCTCCCACTGTGAA 9190
QY 3061 HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer 3080
DB 9191 CACACTTCTTGTGGTCTCTTCCCAATGATACCAATGGCTTCATCAGTGGAGCAGCTCT 9250
QY 3081 TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
DB 9251 TGGAGGAAATGTGATACTTACAGCTGCGAGGCTGAGATATGTCATACAAGGAGCTTCA 9310
QY 3101 AspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeu 3120
DB 9311 GATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCCTTATCCAGTCTGTGAGGCCCTTG 9370
QY 3121 SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr 3140
DB 9371 TCTGTGGGTCCCACTGCTGTGCGCAATGCAATGCGACTGGCACTGGAGAGGACACACCTAT 9430
QY 3141 GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThr 3160
DB 9431 GAAAGTGAAGTGAACCTCAGATGTCTGGAAGTTATACGATGGATACAGATACATACA 9490
QY 3161 PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys 3180
DB 9491 TTCACCTGTGAGAAAGATGGTGGTTCCTGAGAGAAATCTCTCTGAGGAGTCTTAAAAA 9550
QY 3181 CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
DB 9551 TGTCTCTCTCCGGAACATTAACATATATCTTGTTCATGGGACGATTCAGTGTGAAT 9610
QY 3201 ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
DB 9611 AGGCAAGTTCTGTGTCTCATGTGCAGAGAGGTATACCTTTTGGGAGTTAAACATACAGTA 9670
QY 3221 CysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSer 3240
DB 9671 TGTGAGCTTGTATGAACTGGGAGCCACCATTTCTCCGATGAATCTTTCAGTCCAGTTTCT 9730

```
QY 3241 CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
DB 9731 TGTGGAAACCTGAAAGTCAGAAATGATGTTGTGGCAGTAAATACACCTTTGAA 9790
QY 3261 SerThrIleTyrClnCysGluProGlyTyrGluLeuGluClyAsnArgGluArgVal 3280
DB 9791 AGCACAATTTATTTATCAGTGTGAGCTGTGGCTATGAACTAGAGGGGAAACAGGAAACCGTC 9850
QY 3281 CysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGlu 3300
DB 9851 TCCAGGAGAAACAGACAGTGGAGTGGAGGGTGGCAATATGCAAGACAGACAGGTGTGAA 9910
QY 3301 ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsn 3320
DB 9911 ACTCCACTTGAATTTCTCAATGGAAAGCTGACATTTGAAACAGGACGACTGGACCAAC 9970
QY 3321 ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
DB 9971 GTGGTATATTCCTGCAACAGAGCTACAGTCTTGAAGGGCCATCTGAGGACACACTGCAC 10030
QY 3341 GluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValPro 3360
DB 10031 GAAATGGAACTGGAGCCACCCAGTCCCTCTCTGCAACCAATTCATGGCCCTGTTCCT 10090
QY 3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
DB 10091 TTTGTGATTTCCCGAAGTCTCTGCTGTCTGAAAGGAGTTTATGTTGATCAGAAATGAG 10150
QY 3381 SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnPro 3400
DB 10151 TCCATCAATGTAGGGAAGGTTTCTGCTGAGGGCCAGGGCATCATACCTGCAACCCC 10210
QY 3401 AspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAla 3420
DB 10211 GACGAGCTGGACACAGACAGACGCGCAATGTGAAATAATCTCAATGTGTCCACGCT 10270
QY 3421 HisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
DB 10271 CACGTAGAAATGCAATGCTCGAGCGGTACATTATCAATATGAGACATGATCACCTAC 10330
QY 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly 3460
DB 10331 TCATGTACAGTGGATACATGTGGAGGGTTTCTGAGGAGTGTGTTTGTAGAAATGGA 10390
QY 3461 ThrTrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGly 3480
DB 10391 ACATGGACATCACTCTCTATTTGCAGAGCTGTCTGTGATTTCCATGTTCAGATGGGGC 10450
QY 3481 IleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGlu 3500
DB 10451 ATCTGCCAACGCCCAATGCTTCTCTGTCAGAGGGCTGGATGGGGGCGCTCTGTGAA 10510
QY 3501 GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys 3520
DB 10511 GAACCAATCTGCAATCTTCCCTGTGTGAACGGAGGTGCTGTGTGGCCCTTACCACTGT 10570
QY 3521 AspCysProProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
DB 10571 GACTGCCCGCTGGCTGGACGGGTCTCGCTGTATACAGCTGTGTTGCCAGTCTCCCTGC 10630
QY 3541 LeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThrGly 3560
DB 10631 TTAATATGGTGGAAATGTGTAAACCAACCAACCGATGTCACCTGTCTTCTTCTTGGACGGA 10690
QY 3561 HisAsnCysSerArgLysArgThrGlyPhe 3571
DB 10691 CATAACTGTTCCAGGAAAGGAGGACTGGGTTT 10723
RESULT 3
US-10-295-027-429
; Sequence 429, Application US/10295027
; Publication No. US20030232350A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevizi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Ros Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 429
; LENGTH: 10674
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-295-027-429
```

```
Alignment Scores:
Pred. No.: 0 Length: 10674
Score: 19769.00 Matches: 3535
Percent Similarity: 99.41% Conservative: 9
Best Local Similarity: 99.16% Mismatches: 13
Query Match: 98.98% Indels: 8
DB: 16 Gaps: 2
```

US-09-977-053-4 (1-3571) x US-10-295-027-429 (1-10674)

```
QY 1 MetTrpProArgLeuAlaPheCysTyrGlyLeuAlaLeuValSerGlyTyrAlaThr 20
DB 1 ATGTGGCCCTCGCTGGCCCTTTTGTGCTGGGGCTGGCGCTCGCTTCGGGCTGGGCGACC 60
QY 21 PheGlnGluMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
DB 61 TTTCAGCAGATGTCCTCGCTGGCCAAATTTTCAGCTTCGGCTCTTCCCGAGACCGCGGCC 120
QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
DB 121 GGCGCCCCCGGAGTATCCCGCGCGCGCTCTTGGCGAGCAAGCGGGGAGCAGA 180
QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
DB 181 GTGGAGCGCTGGCGCAGCGGTTCGCGCAGCGCTGGCTGCTGCGGAGCTCAGCGAG 240
QY 81 ArgLeuGluLeuValPheLeuValAspSerSerValGlyGluValAsnPheArg 100
```

241 CGCGCTGGAGCTTGTCTTCTGGTGGATGATTCGTCCAGCGTGGCGCAAGTCAACTTCCGC 300
Db
101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
Qy
301 AGCGAGCTCATGTTCTGTCGCAAGCTGCTGTCGACTTTCCTCGTGGTGGCCAGCGGCACG 360
Db
121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
Qy
361 CGCGTGGCCATCTGACCTTCTCGTCCAGAACTACGTGGTGGCGCGTGGATACATC 420
Db
141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAlaIle 160
Qy
421 TCCACCGCGCGCGCGCCAGCAAGTGGCGCTGCTCTCCAGAGATCCCTGCCATC 480
Db
161 SerTyrArgGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeu 180
Qy
481 TCCATCGAGGTGGCGCACCTTACACCAAGGCGCTTCCAGCAAGCGCGCAAAATCTTT 540
Db
181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn 200
Qy
541 CTTTCATGCTAGAGAACTCAACAAAGTTGTATTTCTCATCTACTGATGATTTCCAAT 600
Db
201 GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
Qy
601 GGGGAGACCTTAGACCAATTCAGCGCTCACTGCGAGATTCCAGAGTGGAGATCTTCACT 660
Db
221 PheGlyIleTropGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
Qy
661 TTTGGCATATGGCAAGGAAATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 720
Db
241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260
Qy
721 GAGCACTGTACTCTGCTACACAGTTTGAAGATTTGAGGCTTTAGCTCGCGGCATATG 780
Db
261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
Qy
781 CATGAAGATCTACTTCTGGGAGTTTATTTCAAGATGATATGGTCCACTCTCATATCTT 840
Db
281 CysAspGluGlyAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
Qy
841 TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGGAGCTGCAAAATGGGGAACACACA 900
Db
301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
Qy
901 GGCCATTTGATGTCATCTGAAAGGGGTATACGGGAAGGTCTGCAATGATGATGC 960
Db
321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
Qy
961 ACAGCTTGGCCATCGGGGACATACAAACCTGAAGGCTCACAGGAGGAATCAGCAGTTGC 1020
Db
341 IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys 360
Qy
1021 ATTCCATGTCCTGATGAATAATCACCTCTCCACCTGGGAAGCACATCCCTGAAGACTGT 1080
Db
361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
Qy
1081 GTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGCCACTGCCCTGCC 1140
Db
381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
Qy
1141 CTGAAGCCTCCCGAATAATGTTACTTTATCCAAACACCTTGCAACACCACTTCAATGCA 1200
Db
401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuCys 420
Qy
1201 GCTGTGGGGTCCGATGTACCCCTGGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1260
Db
421 LeuProAsnGlyLeuTyrPheSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
Qy
1261 CTACCCAAATGGTTGTGTCCTGGTTACAGAGACTACTGACAGCTAAGAACATGTCCTCAT 1320
Db
441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
Qy
1321 CTCCCGCCAGCGGAACATGGCCACATCAGCTGTTCTACAAGGGAAATGTTATATAAGACA 1380
Db

461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
Qy
1381 ACATGTTTGGTTCCTGTTGATGAAGGTACAGACTAGAGCAGTGATAGCTTACTTGT 1440
Db
481 GlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
Qy
1441 CAAGGAAACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGGAGCGCCACTGTTCACCC 1500
Db
501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
Qy
1501 TTTTCAGATGCCCAAGATGTCATCATATCCCCCAACTGTGTGCAAGCAGCCAGCCAAA 1560
Db
521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
Qy
1561 TTTTGGAGCATCTGCTATGTAGTTGCGCCAGGGTTTCATTTATCTGGAGTCAAGAA 1620
Db
541 MetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAlaValCysLys 560
Qy
1621 ATGCTGAGATGTACCACTTCTGGAAAATGGAATGTCGAGTTCAAGCAGCTGTGTGTA 1680
Db
561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
Qy
1681 GAGCTGGAGGCTCTCTCAATCAACTGTCTTAAGGACATAGAGGCTAAGACTCTGGAAC 1740
Db
581 GlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
Qy
1741 CAGATTTCTCCATGTTACTTGGCAGATTCCAACACTAAGACAACTCTGTTGANAAG 1800
Db
601 ValSerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAspVal 620
Qy
1801 GTGTGAGTCCACGTTTCATCCAGCTTTTCACTTACCTTTTCCCAATTTGGAGATGTT 1860
Db
621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
Qy
1861 GCTATCGTATACAGCGCACTGACCTTATCGGCAACCAAGCCAGCTGATTTTCCATATC 1920
Db
641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
Qy
1921 AAGTTATTTGATGCAAGAACCACTGTTCATAGACTGGTGCAGATCTCCACCTCCGTC 1980
Db
661 ValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
Qy
1981 GTCTCGAGAGGATACATGCGCAAGCTGGGATGAGCTCAGTCTCAGACAACTCAGCG 2040
Db
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyThr 700
Qy
2041 GCTGNAATTTGGTCAATTTACCAGAGTATACCAAGGAGACCTTTTCCCTCAAGGGGAG 2100
Db
701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Qy
2101 ATAGTACATATACGCCACTGACCCCTCAGGCAATTAACAGGACATGTGATATCCATAT 2160
Db
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
Qy
2161 GTCATAAAAGGTTCTCCCTGTGAAATTTCCATTCACACTGTAAATGGGATTTTATATSC 2220
Db
741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
Qy
2221 ACTCCAGATTAATTTGAGTCAACTGTACATTAATTAATCTTGGAGGGCTATGATTTTCA 2280
Db
761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
Qy
2281 GAAGGTTCTACTGCAAGTATTTATTTGCTTTATGAAGATGGCGTCTGGAAACCAACAT 2340
Db
781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Qy
2341 ACCACTGAATGGCCAGACTGTGCGCAAAAACGTTTTCAAAACCAACGCGGTTCAAGTCT 2400
Db
801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
Qy
2401 GAGATGTTTACAAAGCGCTCGTTGTGTGATGACACAGATCTGATGAAGAAAGTTTCTGAA 2460
Db

QY	821	AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle	840		3520	GGGCGATCTGGAGTTGTTAAAT---		3576
DB	2461	GCAATTTGAGACGACCCCTGGGAAAATGGTCCCATCAATTTGTAGTGATGCAGAGGACATT	2520		1201	ProCysHisAasnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro	1220	
QY	841	AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr	860		3577	CTTGGCCACATAGTGGAAACCTGCCAGCACTTGGGGGTGGTTATGTTGTTCTCTCTCCA	3636	
DB	2521	GACTGCAGACTGGAGGAGAACCTGGACCAAAAATATGTCTAGATATATAATATGACTAT	2580		1221	LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys	1240	
QY	861	GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer	880		3637	CTTGGATATACAGGCTTAAAGTGTGAAAACAGACATCATGATGAGTGCAGCCACTGCTGC	3696	
DB	2581	GAAATGGCTTTGCANTTGGACACGGTGGCTGGGGTGCAGCTAATAGGCTGGATTACTCT	2640		1241	LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly	1260	
QY	881	TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer	900		3697	CTCAACAATGGAGTTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGGCCCTCAGGT	3756	
DB	2641	TACGATGACTTCTCGCACACTGTGCAGAAAACAGCCACAAGCATCGGCAATGCCAAGTCC	2700		1261	TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsn	1280	
QY	901	SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle	920		3757	TACACAGGTCAGCGGTGTGAGAAAATATATAGTGTAGCTCCAGTCTCTGTTTAAAT	3816	
DB	2701	TCACGGATTAAAGAAAGTGGCCCATTTCTGACTATATAAATTAAGTTAATTTTAAACATC	2760		1281	LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal	1300	
QY	921	ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrPduAsnGlnGln	940		3817	AAAGGAATCTGTGTGATGTGTGGCTGCTATCGTTGCACATGTGTGAAGGATTTGTA	3876	
DB	2761	ACAGCTAGTGTGCCATTACCCGATGAGAAATGATACCTTGAATGGGAAAATCAGCA	2820		1301	GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla	1320	
QY	941	ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp	960		3877	GGCCTGCAATTTGAAACAGAAAGTCAATGAATGCCAGTCAACCCATGCTTAAATATGCA	3936	
DB	2821	CGACTCTTCAGACATTGGAACTATACAAATAAACTGAAAGGACCTCTCAACAAAGAC	2880		1321	ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr	1340	
QY	961	ProMetTyrSerPheGluLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu	980		3937	GTCTGTGAAGACACAGGTTGGGGGATTTCTGTGCAAAATGCCACCTGGATTTTGGGTACC	3996	
DB	2881	CCCATGTATTCCTTTGAGTTCGATCAGAAATACCTTATAGCCGACAGCAATTCATTAGAA	2940		1341	ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys	1360	
QY	981	ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal	1000		3997	CGATGTGGAAGAAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGCTACTGT	4056	
DB	2941	ACAAAAGAGGCTTCCCTCTTCGACAGACAGGCTCAGTGTGAGAGGGCGTATGTGTTC	3000		1361	LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys	1380	
QY	1001	AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle	1020		4057	AAAGACGTTGCCAATAGCTTCAGATGCCCTGTGTGAGCTGGCTTTCACAGATCACACTGT	4116	
DB	3001	AAATGCCCTTTGGGAACCTTATATAATCTGGAACATTTTCACCTGTGAAAGCTGCCGGATC	3060		1381	GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp	1400	
QY	1021	GlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr	1040		4117	GAATTTGAACATCAATGAATGTCACTTAATCCATGTAGAAATCAGGCCACCTGTGTGGAT	4176	
DB	3061	GGATCCTATCAAGATGAAGAAGGCACTTGATGTGAAGCTTTTGCCCTCTGGGATGTAC	3120		1401	GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr	1420	
QY	1041	ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr	1060		4177	GAATTTAAATTCATACAGTTGTAAATGTACGAGGATTTTTCAGGCAAAAAGGTGTGAACA	4236	
DB	3121	ACGGAATATATCCATTCAGAAACATCTCTGATTGTAAAGCTCAGTGTAAACAAGGCACC	3180		1421	GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet	1440	
QY	1061	TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys	1080		4237	GAACAGTCTACAGGCTTTAACTGGATTTTGAGTTTCTGGCATCTATGGATATGTCATG	4296	
DB	3181	TACTCATACAGTGGACTTGGAGACTTGTGAATCGTGTCCACTGGGGCACTTATCAGCCAAA	3240		1441	LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer	1460	
QY	1081	PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla	1100		4297	CTAGATGGCATGCTCCCATCTCTCCATGCTTAACCTGTACCTTCTGTGATGAAATCCTCT	4356	
DB	3241	TTTGGTTCCCGGAGCTGCTCTCTGTGTCAGAAACACCTCAACTGTGAAAGAAGAGGCC	3300		1461	AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr	1480	
QY	1101	ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu	1120		4357	GACGCATGAACATATGGAACACCAATCTCTATGAGTTGTATACCGCGAGCGACAATACC	4416	
DB	3301	GTGAACATTTCTGCATGTGGAGTTCCTGTCTCAGAAAGGAAATTCCTGGGTTCCTGGTTA	3360		1481	LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr	1500	
QY	1121	MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys	1140		4417	TTGCTCTTGACATGATTAACCGCTGGGTCTTTATGTGATGGCAGGAGGAAAGATACAA	4476	
DB	3361	ATGCCCTGTACCCCATGCTCTCGTGACTATTACCAACCTTAATGCGAGGAGGCCCTTCTGC	3420		1501	AsnCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAla	1520	
QY	1141	LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys	1160		4477	AACTGTCTCTCGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGACAGTGCC	4536	
DB	3421	CTGGCTGTGCTTTTATGGAATCACTCCCATTTGCTGTGTTCCAGATTCATCAAGATGT	3480		1521	AsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer	1540	
QY	1161	SerSerPheSerThrThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu	1180		4537	AATGCGATCTGGAAGTCTATATCGATGGGAAATATATCTGACGGTGGTGTGCGCTCTCT	4596	
DB	3481	TCACT-----TCAGTTCTGAATATTACTATTTTTCGTTGGATTT	3519		1541	ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLys	1560	
QY	1181	GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn	1200					

4597	Db	GTGGTTTGCCCATPACTCGTGGTGGTGGCTAGTCTTCGGGGCAAGACAGACAAAAA	4655
1561	Qy	GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp	1580
4657	Db	GGAGNGGGATTACGCCCNAGCTGAGCTCTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG	4716
1581	Qy	AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu	1600
4717	Db	GACTATGTCCTGCTCCACACAGCAGTGAAGTCACTGGCTACTCTCTGCCACAGGAACCTC	4776
1601	Qy	SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys	1620
4777	Db	AGTAAAGGAACGTTGATGACATGCCCTGATTTCTTTGTTCAGGAATTTGGGGGAAGTGAAG	4836
1621	Qy	IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis	1640
4837	Db	ATCGATTCTAAGACCATATTTTGTCTGATTTGCCACGCTTAGGAGGGTCACTGCGCTCAT	4896
1641	Qy	LeuArgThrAlaSerGluAspLeuLeuProGlySerLysValAsnLeuPheCysAspPro	1660
4897	Db	CTGAGAACTGCATCTGAAGATTTTAAAGCCAGGTTCCAAAGTCAATCTGTTCTGTGATCCA	4956
1661	Qy	GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln	1680
4957	Db	GGCTTCAGCTGGTGGGAAACCCCTGTGCAGTACTGTCTGAATCAAGACACAGTGGACACA	5016
1681	Qy	ProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPhe	1700
5017	Db	CCACTTCTCACTGTGAACGCATTTAGCTGTGGGGTGCCACTCTCTTTGGAGATGGCTTC	5076
1701	Qy	HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr	1720
5077	Db	CATTCAGCCGATGACTTCTATGCTGGCAGCACAGTAACCTACCACTGTCACCAATGGCTAC	5136
1721	Qy	TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer	1740
5137	Db	TATCTATTGGTGACTCAAGGATGTTCTGTACAGATAATGGAGCTGGAAACGGCGTTTCA	5196
1741	Qy	ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer	1760
5197	Db	CCATCTCGCCCTTGATGTCGATGAGTGTGCAGTTGGATCGAGTTGTAGTGAGCATGCTTCT	5256
1761	Qy	CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly	1780
5257	Db	TGCTCGAAGCTAGATGATCCATCATATGTTCTATGTTGCCACCGTACACAGGAGATGGG	5316
1781	Qy	LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer	1800
5317	Db	AAAAACGTGCAGAACCTATAAAATGTAAGGCTCCGGAATCCGGAATAATGGCCATCC	5376
1801	Qy	SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln	1820
5377	Db	TCAGTGCAGATTATACATGATGTTGCCGAGTCAATTTTCTGTCTAGGAAGGATACAG	5436
1821	Qy	LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro	1840
5437	Db	TTGATGGGAGTAACCAAAATCACATGTTTGGAGTCTGCAGAAATGCAATCATCTAAATACCA	5496
1841	Qy	TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu	1860
5497	Db	TATTGTAAGCTGTTTTCATGTGGTAAACCGGCTATTTCGAGAAATGGTTGCATTGGAGG	5556
1861	Qy	LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla	1880
5557	Db	TTAGCATTTTACTTTTGGCAGCAAGTGACATATAGGTGTATATAAGGATATATCTCTGCC	5616
1881	Qy	GlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCys	1900
5617	Db	GGTGATAAAGAAATCATCTCTGTCTGCTAACAGTTTCTTGGAGTCAATCCCTCCTGTGTGT	5676
1901	Qy	GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly	1920
5677	Db	GAACCACTGAAGTCTTCTAGTCCGAAAAATATAATATCGAAAAATATATTTTGAGTGGG	5736

Qy	1921	LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro	194
Db	5737	CTTACCTACCTTTCTACTGTCATCATATTCATGGCATACAGATACAGCTTTACAGGGCCCT	5796
Qy	1941	SerIleIleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHisLeu	1960
Db	5797	TCCATATTGAATGCACGGCTCTTGGCANTCTGGACAGAGCGCACCTGCTGTCACTTC	5856
Qy	1961	ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr	1980
Db	5857	GTCTCTCTGGAGAACCACTGCCTCAATCAAGATGCTGTCATTACGGGGAATAAATCTCACT	5916
Qy	1981	PheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr	2000
Db	5917	TTTCAGGAACACCGTCACTTACACTTGTCAAGAAGGCTATATACTCTTGCTGTGCTTGAACCC	5976
Qy	2001	IleGluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaValSer	2020
Db	5977	ATTGAATGCCCTGGCCGAGGCAAGTGGAGTAGAAGTGACCCAGCAGTGCCTGGCTGTCTCC	6036
Qy	2021	CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly	2040
Db	6037	TGTGATGAGCACCCCATTTGTGGACCAAGCCCTCTCCAGAGACTGCCCATCGGCTCTTTGGA	6096
Qy	2041	AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu	2060
Db	6097	GACATTGCATTTACTACTGCTCTGATGGTTTACGCTAGCAGACAAATTCAGCTTCTC	6156
Qy	2061	CysAsnAlaGlnGlyTyrIleValProProGluGlyGlnAspMetProArgCysIleAla	2080
Db	6157	TGCAATGCCCAGGGCAAGTGGGTACCCCAAGAGGTCACAGACATGCCCGCTTGTATAGCT	6216
Qy	2081	HisPheCysGlnLysProProSerValSerTyrSerIleLeuGluSerValSerLysAla	2100
Db	6217	CATTTCCTGTAAGAAACCTCCATCGGTTTCTATAGCATCTTTGGAACTCTGTGAGCAAGCA	6276
Qy	2101	LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr	2120
Db	6277	AAATTTCCAGCTGGCTCAGTTGTGAGCTTTAAATGCATGGAAGGCTTTGTACTGAACACC	6336
Qy	2121	SerAlaLysIleGluCysMetArgGlyGlyGlnTyrAsnProSerProMetSerIleGln	2140
Db	6337	TCAGCAAGAATTCGAATGATGAGAGGTGGCAGTGGAAACCTTTCCCCCATGTCCATCCAG	6396
Qy	2141	CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer	2160
Db	6397	TGCATCCTGTGCGGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCA	6456
Qy	2161	AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly	2180
Db	6457	AACTACAGTTTTTGAGCCATGGTGGCTTACAGCTGCAACAAGGGGTCTTACATCAAAAGG	6516
Qy	2181	GlnLysLysSerThrCysGluAlaThrGlyGlnTyrSerSerProIleProThrCysHis	2200
Db	6517	GAAAGAAGAGACACTTCGGAAGCCACAGGGCAGTGGAGTAGTCTTATACCAGCAGTGGCAC	6576
Qy	2201	ProValSerCysGlyGluProProLysValIleAsnGlyPheLeuGluHisThrThrGly	2220
Db	6577	CCGGTATCTTGTGGTGAACCCACCTAAGTTTGAGATGGCTTTCTGCGAGCATACAACCTGGC	6636
Qy	2221	ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer	2240
Db	6637	AGGATCTTTGAGAGTGAAGTGAAGGTATCAGTGTAAACCCGGGCTTATAGTCAGTTCGGAAGT	6696
Qy	2241	ProValPheValCysGlnAlaAsnArgHisTyrHisSerGluSerProLeuMetCysVal	2260
Db	6697	CCTGTATTATTGTCTGCCCAAGCCATATGCCACTGGCACAGTAGAATCCCTCTCGATGTGTGT	6756
Qy	2261	ProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe	2280
Db	6757	CCTCTGCAGCTGTGGAACCTCCCGGCATCCAGATGGCTTCTATGAAGAGAGAAAACTTT	6816

Qy 2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
Db 6817 GAAGTAGGTCCTCAAGGTTTCAGTTTTCTGTAAATAGGGTTATGACCTTTGTTGGTGACAGT 6876
Qy 2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysSerAsnProLysCysMetPro 2320
Db 6877 TCTTGGACATGTGAGAAATCTGGCAANTGGATAGAAAGTCAATCCAAAGTGCATGCT 6936
Qy 2321 AlalysCysProGluProProLeuLeuGlnLeuValLeuLysGluLeuThrThr 2340
Db 6937 GCCAAGTCCCGAGAGCCGCCCTCTGGAAACACAGCTAGTATTAAGAGAGTTGACCAACC 6996
Qy 2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
Db 6997 GAGGTAGAGTGTGACATTTTCTGTAAAGAGGGCATGCTCCGAAGGCCCTCTCTGTC 7056
Qy 2361 LeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
Db 7057 CTGAANTGCTTGCCATCCAGCAATGGAAATGACTCTTCCCTGTTTGTAAAGATTGTTCTT 7116
Qy 2381 CysThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
Db 7117 TGTACCCACCTCCCTAAATTTCTTGTGTGCTCCCAATTCCTTCTCTCTCTCTCTCTCT 7176
Qy 2401 GlySerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThr 2420
Db 7177 GGAAGTACTGTCAAGTATTTCTTGTGTAGTGGTGTTCCTTAAGAGGAAATTTCTATCCACC 7236
Qy 2421 LeuCysGlnProAspGlyThrTrpSerProLeuProGluCysValProValGluCys 2440
Db 7237 CTCTGCCAACTGATGGACCTGGAGCTCTCACTGCCAGAAATGTTCCAGTAGAATGT 7296
Qy 2441 ProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlnGlyLeuAlaTyrLeuSer 2460
Db 7297 CCCCAACCTGAGAAATCCCAATGGAATCAATGATGTGCAAGGCTTGCCTATCTCAGC 7356
Qy 2461 ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys 2480
Db 7357 ACAGCTCTCTATACCTGCAAGCCAGGCTTTGAATGGTGGGAAATACTACCAACCTTTGT 7416
Qy 2481 GlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLys 2500
Db 7417 CGAAGAAATGGTCACTGGCTGGAGGAAACCAACATGTAAGCCATTGAGTGCCTGAAA 7476
Qy 2501 ProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal 2520
Db 7477 CCCAAGGAGATTTGAAATGGCAATTTCTTACACGGACCTACACTATGACACAGCCGTT 7536
Qy 2521 ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu 2540
Db 7537 ACCTTACTCTTGCAACCGAGGCTTTTGGCTCGAAGGTCCTGAGGCTTGACCTGTTTAGAG 7596
Qy 2541 ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln 2560
Db 7597 ACAGGTGATGGAGTAGATGCCCATCTTGCATGCCATCCATGTCATGTCATGCCACAA 7656
Qy 2561 ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyr 2580
Db 7657 CCCATTTGAAATGGTTTTGTAGAAGGTGCAGATTACAGCTATGGTGCCATAATCATCTAC 7716
Qy 2581 SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly 2600
Db 7717 AGTTGCTCTCCCTGGGTTTCAGGTGGCTGTGCATGCCATGCCAGCTGTGAGAGATCAGA 7776
Qy 2601 TrpSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp 2620
Db 7777 TGTCAAGTTCCATCCCAACATGATGCCAATAGACTGTGGCTCCCTCTCTCATATAGAT 7836
Qy 2621 PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspMet 2640
Db 7837 TTTGGAGACTGTACTAAATCAAAAGATGACCGAGATATTTTGGACAAAGACGACATG 7896
Qy 2641 MetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThr 2660

Db 7897 ATGGAGTTCCATATGTGACCTCTCCACCTCTTATATCATTTGGAGAGTGCTAAAC 7956
Qy 2661 TrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet 2680
Db 7957 TGGGAAATATACAAAGAGTCTCTGTCTACACATTCAACAACTTTCTGTATGTATGACCATG 8016
Qy 2681 ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGln 2700
Db 8017 GTTTCATACACTGTAATCCAGGATATGAATCTTGGGGAACCCCTGTGCTGATGTGCCAG 8076
Qy 2701 GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
Db 8077 GAAGATGACACTTGAATGGCAGTCACCATCTCTGCATTTCAATTTGAATGTGACTTGCCT 8136
Qy 2721 ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln 2740
Db 8137 ACTGCTCTGAAATGGCTTTTGTGCTTTTACAGAGACTAGCATGGGAAGTCTGTGTCAG 8196
Qy 2741 TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
Db 8197 TATAGCTGTAAACCTGGACACATTTCTAGAGGCTCTGACTTAAGGCTTTGTCTAGGAAT 8256
Qy 2761 ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
Db 8257 AGAAAGTGGAGTGTGCTCTCCACGCTGTGAGGCATTTCAATGCAAAAGCCAAATCCA 8316
Qy 2781 ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu 2800
Db 8317 GTCATGAATGGATCCTCAAGGAAGCAACTACACATACCTGAGCAGCTGTACTATGAG 8376
Qy 2801 CysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn 2820
Db 8377 TGTGACCCCGGATATGTCTGAATGGCACTGAGAGAGAACATGCGAGATGACAAAAC 8436
Qy 2821 TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla 2840
Db 8437 TGGGATGAGGATGAGCCCATTTGCATTTCTGTGGACTGCGAGTTCACTCCAGCTCTCAGCC 8496
Qy 2841 AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn 2860
Db 8497 AATGGCCAGGTGAGAGAGAGACGAGTACATTTCCAAAAGAGATTGAATACACTTGCAT 8556
Qy 2861 GluGlyPheLeuLeuGluGlyValaArgSerArgValCysLeuAlaAsnGlySerTrpSer 2880
Db 8557 GAAGGTTCTTCTCTGAGGAGCCAGGAGTCCGGTTTGTCTTGCCAAATGGAGTTGGAGT 8616
Qy 2881 GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly 2900
Db 8617 GGAGCCACTCCCGACTGTGTGCTGTGCAGATGTGCCACCCGCCCACTGGCCAAATGGG 8676
Qy 2901 ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly 2920
Db 8677 GTGACGAGGCTCTGGACTATGGCTTTCATGAAGGAAGTAACATTCCTGTCACGAGGCG 8736
Qy 2921 TyrIleLeuHisGlyValaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu 2940
Db 8737 TACATCTTGACGGTGTCTCAAAACTCACTGTCTAGTCAGATGCCAACTGGGATGACAG 8796
Qy 2941 IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
Db 8797 ATTCTCTCTGTAAACCCAGTCAACTGTGGACCTCTCTGAAGATCTTGCCCAATGGTTTCCCT 8856
Qy 2961 AsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980
Db 8857 AATGGTTTTTCTTATTCATGCGGGCCATATACAGTATCAGTGTCTTCTGTTTAAAG 8916
Qy 2981 LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerPro 3000
Db 8917 CTCCTGGAATTCATCAGAGAGGTCCTCTCTCATGGCTCTCTGAGTGGCAGCTCACCT 8976
Qy 3001 SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr 3020

Db 8977 TCCTGCTGCTTGCAGATGTTCCACACCAAGTAATTCGAATATGGAACCTGTCATATGGGACA 9036
Qy 3021 AspPheAspCysGlyLysAlaIleAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
Db 9037 GATTTTGTACTGTGGAAAGGAGCCGCGGATTCAGTGTCTTCAAGGCTTCAAGCTTCCTAGGA 9096
Qy 3041 LeuSerGluIleThrCysGlyLysAlaAspGlyGlnTyrSerSerGlyPheProHisCysGlu 3060
Db 9097 CTTTCTGGAATACCTGTGAGCGATGGCCAGTGGAGCTCTGGGTTCCTCCCTCAGTGTGAA 9156
Qy 3061 HisThrSerCysGlySerLeuProMetIleProAsnAlaPheLeuSerGluThrSerSer 3080
Db 9157 CACACTTCTTGTGTCTCTCTTCCAAATGATACCAATGCGTTCATCAGTCAGACCGAGCTCT 9216
Qy 3081 TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
Db 9217 TGGAAAGAAATGTGATAACTTACAGCTGAGGTCTGGATATGTATACAAAGGAGGTTCAC 9276
Qy 3101 AspLeuIleCysThrGluLysGlyValTyrSerGlnProTyrProValCysGluProLeu 3120
Db 9277 GATCTGATTTGTACAGAAAGGGGTATGGCCAGCTTATCCAGTCTGTGAGCCCTTG 9336
Qy 3121 SerCysGlySerProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr 3140
Db 9337 TCCTGTGGGTCCCCACCGTCTGTGCGCAATGTCAGTGGCAACTGGAGAGGCACACACCTAT 9396
Qy 3141 GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThr 3160
Db 9397 GAAAGTGAAGTGAACACTCAGATGTCTGGAAGTTATACGATGGATACAGATACAGATACA 9456
Qy 3161 PheThrCysGlnLysAspGlyArgTyrPheProGluArgIleSerCysSerProLysLys 3180
Db 9457 TTCACCTGTGAGAAAGATGTGTGTCTCTGAGAGATCTCTCTGAGTCTCTTAAAAA 9516
Qy 3181 CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
Db 9517 TGTCTCTCCGGAACATACATATATCTGTACATGGGACGATTTTCAGTGTGAAT 9576
Qy 3201 ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
Db 9577 AGGCAAGTTTCTGTGTCTGTGTCAGAGAGGGTATACCTTTGAGGAGTTAAATATCAGTA 9636
Qy 3221 CysGlnLeuAspGlyThrTyrGluProPheSerAspGluSerCysSerProValSer 3240
Db 9637 GTGCAGCTTGATGGAACCTGGAGGCCACCACTTCTCCATGATCTTCAGTCTCAGTTCT 9696
Qy 3241 CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
Db 9697 TGTGGGAAACCTGAAAGTCCAGAACATGGATTTGTGTGTGGTGGCAGTAAATACACCTTTGAA 9756
Qy 3261 SerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgVal 3280
Db 9757 AGCACATTTATTTATCATGTGTGAGCCCTGGCTGTATGACTAGAGGGGGAACAGGAAACGTGTC 9816
Qy 3281 CysGlnGluAsnArgGlnTyrSerGlyValAlaIleCysLysGluThrArgCysGlu 3300
Db 9817 TGCAGAGAAACAGACAGTGGATGGAGGGGTGCAATATGCAAGAGACAGAGTGTGAA 9876
Qy 3301 ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsn 3320
Db 9877 ACTCCACTTGAATTTCTCAATGGGAAAGCTGACATTTGAAACAGAGAGACTGGACCCCAAC 9936
Qy 3321 ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
Db 9937 GTGGTATTTCTTCGCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCACACTGCACA 9996
Qy 3341 GluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValPro 3360
Db 9997 GAAATGGAACTGGAGCCACCCAGTCCCTCTCTGCAAAACCAATCATTCCTGCTTCTCT 10056
Qy 3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
Db 10057 TTTGTGATTTCCCGAGATGCTCTGCTGTCTGAAAGGAGTTTTATGTGTGATCAGATGTG 10116

Qy 3381 SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnPro 3400
Db 10117 TCCATCAATGTAGGAGAGGTTCCTGTCGAGGGCCACGGCATCATTCACCTGCAACCCC 10176
Qy 3401 AspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAla 3420
Db 10177 GACGAGACGTGGACACACAGACAGCCCAATGTGAAAAATCTCATGTGTGTCACACAGCT 10236
Qy 3421 HisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
Db 10237 CACGTAGAAAATGCAATTTGCTCGAGGGGTACATTTATCAATATGGAGACATGATCACCTAC 10296
Qy 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly 3460
Db 10297 TCATGTACAGTGGATACATGTTGGAGGTTTCTCTGAGGAGTGTGTTTGTAGAAAATGGA 10356
Qy 3461 ThrTrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGly 3480
Db 10357 ACATGGACATCACCTCTTATTTGACAGAGCTGTCTGTGCAATTTCCATGTGCAGAAATGGGGC 10416
Qy 3481 IleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGlu 3500
Db 10417 ATCTGCCAACCGCCCAATGCTTGTCTGTCTGACAGGGCTGATGGGGCGCTCTGTGAA 10476
Qy 3501 GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys 3520
Db 10477 GAACCAATCTGCATCTTCTCTGTCGACGAGGTGCTGTGTGGCCCTTACCAGTGT 10536
Qy 3521 AspCysProGlyTyrThrThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
Db 10537 GACTGCCGCTGCTGCTGACGGGCTCTCGCTGTATACAGCTGTGTGTCAGTCTCCCTGC 10596
Qy 3541 LeuAsnGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThrGly 3560
Db 10597 TTAATGCTGGAATAATGTGAAGACCAACCGATGCTACTGCTCTTCTTCTTGGACGGGA 10656
Qy 3561 HisAsnCysSerArg 3565
Db 10657 CATNACTGTTCAGG 10671

RESULT 4
US-10-295-027-1080
; Sequence 1080, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14

```

; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1080
; LENGTH: 10674
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-1080

Alignment Scores:
Pred. No.: 0 Length: 10674
Score: 19769.00 Matches: 3535
Percent Similarity: 99.41% Conservative: 9
Best Local Similarity: 99.16% Mismatches: 13
Query Match: 98.98% Indels: 8
DB: 16 Gaps: 2

US-09-977-053-4 (1-3571) x US-10-295-027-1080 (1-10674)
QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
DB 1 ATGTGGCCCTCGCTGGCCCTTTTGTGCTGGGCTCGGGCTCGCTGTTTCGGGCTGGGGCACC 60
QY 21 PheGlnGluMetSerProSerArgAnnPheSerPheArgLeuPheProGluThrAlaPro 40
DB 61 TTTCAGCAGATGTCCCGCTGCGGCAATTTTCAGCTTCCTCCGCTCTTCCTCCGAGACCGGCGCC 120
QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
DB 121 GGGGCCCCGGGAGATATCCCGGCGCCCGCTCTCTGGGACGAGACGGGGGAGCAGA 180
QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
DB 181 GTGGAGCGGCTGGGCGAGCGTTCGCGACGCGTGGCTGGCTGGGCGAGCTCAGCGAG 240
QY 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAlaPheArg 100
DB 241 CCGCTGGAGCTTGTCTCTCTGTGGTGGATGATTCGTCCAGCGTGGCGGCAAGTCAATTCGCG 300
QY 101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
DB 301 ACGGAGCTCATGTTCTGTCGCGAGCTGCTGTCCGACTTCCCGGTGGTGGCCCGGCGCAG 360
QY 121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
DB 361 CGCGTGGCCATCGTGACCTTCTCGTCCAGAACTACGTGTGCGCGCGCTCGATTCATATC 420
QY 141 SerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGluGluLeuProAlaIle 160
DB 421 TCCACCGCGCGCGCGCGCAGCAAGTGGCGGCTCTCTCCAGAGATCCCTGCGCATC 480
QY 161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyValAlaPheGlnAlaAlaGlnIleLeu 180
DB 481 TCTACCGAGGTGGCGGCACTTACACCAAGGGCGCTTCCAGCAGCGCGGCAATTCCTT 540
QY 181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn 200
DB 541 CTTTCATGTAGAGAACTCAACAAAGTTGATTTCTCATCTCATCTGATGGATATTCAT 600
QY 201 GlyLysAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
DB 601 GGGGAGAGCCCTAGACCAATTCAGCGCTCAGTGGGATTCAGGATGAGATCTTCAT 660
QY 221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240

```

```

DB 661 TTTGGCATATGCAAGGGAACATTTCGAGAGCTGAATGATGATGCTTCCACCCCAAGGAG 720
QY 241 GluHisCysTyrIleuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeu 260
DB 721 GAGCACTGTTACCTGCTACACAGATTTTGAGGAATTTTGAGGCTTTAGCTCGCGGCAATG 780
QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
DB 781 CATGAGATCTACCTTCTCGGAGTTTATTCAGATGATATGCTCCACTGCTCATATCTT 840
QY 281 CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
DB 841 TGTGATGAAGGCAAGCACTGCTGTGACCGAATGGGAAGCTGCAATGTGGGACACACACA 900
QY 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
DB 901 GGCCATTTTGGAGTGCATCTGTGAAAGGGGTATTACGGGAAAGGCTGCGAGTATGAATGC 960
QY 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
DB 961 ACGCTTGGCCATCGGGGACATCAAACTGAAGGCTCACCAGAGGGAATCAGCAGTTGC 1020
QY 341 IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys 360
DB 1021 ATTCCATGCTCTGATGAAATCACACTCTCCACCTGGAAGCACATCCCTGGAAGCTGT 1080
QY 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
DB 1081 GTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGCCATCGCCCTGCC 1140
QY 381 LeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
DB 1141 CTGAGCCCTCCCGAAATGTTACTTTATCCAAACACTTGCACCAACCACTTCATATGCA 1200
QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuCys 420
DB 1201 GCTGTGGGCTCGATGTCACTGATTTGATCTGTGGGAGGAGCAGCATCATCTATGT 1260
QY 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
DB 1261 CTACCAATATGTTGTGTGTCGCGTTCAGAGAGCTACTGCAGAGTAAGAACTGTGCTCAT 1320
QY 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1321 CTCGCGCACCGGAACATGCGCCACATCAGCTGTCTCAAGGGGAATGTTATATAGACA 1380
QY 461 ThrCysLeuValAlaCysAspGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1381 ACATGTTTGGTTCCTGTGATGAAGGCTACAGACTAGAGGCACTGATAGCTTACTTGT 1440
QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1441 CAAGGAAACAGCGCGGATGGGATGGCGAGAACCCCGTGTGTGGAGCGCCTGTTCACCC 1500
QY 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
DB 1501 TTTTCAGTCCCAAGATGTCATATATCCCCCACACACTGTGGCAGCAGCAGCCAGCAA 1560
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1561 TTTGGAGCATCTGCTATGTAAGTTCGCCCAAGGGTTCATTTATCTGGAGTCAAGAA 1620
QY 541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaIleValCysLys 560
DB 1621 ATGCTGAGATGTACCACTTCTGGAAATGGAATGTGCGAGTTCAGGAGCTGTGTGTAA 1680
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
DB 1681 GACGTGGAGCTCTCTCAATCAACTGTCTTAGGACATAGAGCTTAAGACTCTGGRACAG 1740
QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 1741 CAAGATTCTGCCAATGTTTACCTGGCAGATTCCAACAGCTAAAGACCACTCTGCTGTAAG 1800

```

601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
1801 GTGTAGTCCAGTTCATCCAGCTTCACCCACCTTACTCTTTCCAAATTGAGATGT 1860
621 AlaIleValTyrThrAlaThrAspLeuSerGlyValGlnAlaSerCysIlePheHisIle 640
1861 GCTATCGTATACAGGCAACTGACCTATCCGGCAACAGCCAGCTGCATTTTCCATATC 1920
641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
1921 AAGGTTATTGATGCGAACCCACCTGTATAGACTGGTGCGAGATCTCCACCTCCCGTCCAG 1980
661 ValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
1981 GTCTCGGAGAGGTACATGCCGCACTGGATGAGCTCTGATCTCTGAGCACTCAGGG 2040
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
2041 GCTGAATTGGTCAATACCAAGAGTCATACAAAGGAGACCTTTTCCCTCAAGGGGAGACT 2100
701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
2101 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGTATATCCATATT 2160
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
2161 GTCATAAAGGTCTCCCTGTGAATTCATTCACACCTGTAAATGGGATTTTATATGC 2220
741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
2221 ACTCCAGATAATACTGGAGTCAACTGTACATTAACTTGCTTGGAGGGCTATGATTTTCA 2280
761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
2281 GAAGGGTCTACTGACCAAGTATTATTGTGCTTATGAAGATGGCGTCTGGAACCAACATAT 2340
781 ThrThrGluTyrProAspCysValAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
2341 ACCACTGAATGGCCAGACTGTGCAAAAACGTTTTCAAACCAACGGGTTCAAGTCTTTT 2400
801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
2401 GAGATGTTCTACAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAAGTTTCTGAA 2460
821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
2461 GCATTTGAGACGACCTCGGCAAAATGGTCCCATCATTTTGTAGTATGACAGAGACATT 2520
841 AspCysArgGluLeuGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
2521 GACTGCAGACTGGAGGAGAACCTGACCAAAAATATTGGCTAGATAATAATTATGACTAT 2580
861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrPheGlyAlaAlaAsnArgLeuAspTyrSer 880
2581 GAAATGGCTTTGCAATTGGACCGAGTGGCTGGGGTGCAGCTAATAGCTGGATTACTCT 2640
881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
2641 TAGCATGACTTCTCGACACTGTGCAAGAAACAGCCCAAGCATCGGCAATGCCAAGTCC 2700
901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
2701 TCACGGATTAAGAAGTGGCCCATTTATCTGACTATAAATAATTAATTTTAACTC 2760
921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGln 940
2761 ACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAATCAGCAA 2820
941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
2821 CGACTCTTCAGACATTGGAACTATCACAAATAAACTGAAGAGGACTCTCAACAAAGAC 2880

961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
2881 CCCATGTTATCTTTTTCAGCTTTCATCAGAAATACTTATAGCCGACAGCAATTCATTAGAA 2940
981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
2941 ACAAAGAAGCTTCCCTCTTCAGACCAAGGCTCAGTGTCTGAGAGGGCGTATGTGTCTC 3000
1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
3001 AATTGCCCTTTGGGAACCTATTATAATCTGGAACATTTTCACTGTGAAGCTGCCGGATC 3060
1021 GlySerTyrGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
3061 GGATCTATCAAGATGAAGAAGGCAACTTGGTGCAGCTTTTGCCTCTCTGGGATGTAC 3120
1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
3121 ACGGAATATATCCATTCAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAAGGACACC 3180
1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
3181 TACTCATACAGTGGACTTGGAGCTTGTGAATCGTGTCCACTTGGGCACCTTATCAGCCAAA 3240
1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
3241 TTTGGTTCCTCGGAGCTGCTCTCGTGTCCAGAAAGCAACCTCAACTGTGAAGAAGGAGGCC 3300
1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
3301 GTCAACATTTCTGCATGTGGAGTTCCTTGTCCAGAAAGCAAAATCTCGCGTTCCTGGGTTA 3360
1121 MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
3361 ATGCGCTGTCAACCATGCTCTCGTGACTTATCCACTAATGATGAGGGAAGCCCTTCTGCG 3420
1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
3421 CTGGCTGTCTCTTTTATGGAACCTACCCCATTCGCTGCTGCTCCAGATCCATCACAGATGT 3480
1161 SerSerPheSerSerThrPheSerAlaAlaGluSerValValProProAlaSerLeu 1180
3481 TCAACT-----TCAGTTCTGAATATTACTATTTCCTCGTGAGATT 3519
1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
3520 GGGCATCTGAGTGTGTTAAAT---TGTCTTCTGAGGTTTTCCATGAATGCTTCTTTAAC 3576
1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
3577 CTTGCCACAAATAGTGGAACTTGGCCAGCACTTGGCGTGGTTATGTTTGTCTCTGTCCA 3636
1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
3637 CTTGGATATACAGCTTAAAGTGTGAACACACATCGATGAGTGCAGCCCATGCTTTCG 3696
1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
3697 CTCACAAATGGAGTTGTAAAGACCTAGTTGGGGAATTCATTGTGAGTGGCCATCAGT 3756
1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
3757 TACACAGCTCAGCGTGTGAAGAAATATAAATCAGTGTAGTCTCCAGTCTCTGTTTAAAT 3816
1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
3817 AAAGGAATCTGTGTGTGTGTGGTGGCTGCTATCGTTCACATGTGTGTAAGAGATTGTA 3876
1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
3877 GGCTGATGTGTGAACAGAGATCTAATGAATGCCAGTCAAAACCCATGCTTAAATATGCA 3936
1321 ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340

Db	3937	GTCTGTGAAGACAGGCTGGGGGATTTCTGTGCAAAATGCCACCTCGGATTTTGGGTACC	3996
Qy	1341	ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys	1360
Db	3997	CGATGTGGAAAGACGTGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGCTACCTGT	4056
Qy	1361	LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys	1380
Db	4057	AAAGACGGTGGCAATAGCTTCAGATCCCTGTGTGTCAGCTGGCTTCACAGGATCACACTGT	4116
Qy	1381	GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp	1400
Db	4117	GAATTTGAACATCAATGAATGTGAGTCTAATCCATGTAGAAATCAGGCCACCTGTGTGGAT	4176
Qy	1401	GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr	1420
Db	4177	GAATTTAAATTCATACATTTGTAATGTACGCCAGGATTTTCAGGCCAAAGGTGTGAAACA	4236
Qy	1421	GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet	1440
Db	4237	GAACAGCTACAGGCTTTAAACCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTGATG	4296
Qy	1441	LeuAspGlyMetLeuProSerIleuHisAlaLeuThrCysThrPheThrMetLysSerSer	1460
Db	4297	CTAGATGGCATGCTGCCATCTCTCCATGCTCTAAACCTGTACCTTCTGGATGAATCTCT	4356
Qy	1461	AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr	1480
Db	4357	GACGACATGAACATATGGAACACCAATCTCTATGCGAGTTGTAAACGGCAGCAGATACC	4416
Qy	1481	LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr	1500
Db	4417	TTGCTCTGACTGATTTATAACGGCTGGGTCTTTTATGTGAATGGCAGGCAAGATAACA	4476
Qy	1501	AsnCysProSerValAsnAspGlyArgThrHisIleAlaIleThrTrpThrSerAla	1520
Db	4477	AACTGTGCTCCCTGGAATGATGGCAGATGGCATCATATTGCAATCATCTTGGACAAGTGCC	4536
Qy	1521	AsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyAlaGlyLeuSer	1540
Db	4537	AATGGCATCTGGAAAGTCTATATCGATGGGAAATATATCTGACGGTGGTGTCTGGCTCTCT	4596
Qy	1541	ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysLys	1560
Db	4597	GTGTGTTTGGCCATACCTGCTGGTGGTGGCTTAGTTCTGGGGCAAGCAGCAAGCAAAAA	4656
Qy	1561	GlyGluGlyPheSerProIleGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp	1580
Db	4657	GGAGAGGGGATTCAGCCCACTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG	4716
Qy	1581	AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu	1600
Db	4717	GACTATGCTCTGTCTCCACAGCAGGTGAAGTACTGCTGCTCTGCCCCAGAGGAACTC	4776
Qy	1601	SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys	1620
Db	4777	AGTAAAGGAAACGTGTGTAGCATGGCTGTATTTCTGTGAGGAATGTGGGAAAGTGAAG	4836
Qy	1621	IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis	1640
Db	4837	ATCCGATTTCTAAAGACATATTTTGTCTGATTTGCCACCGCTTAGGAGGGGTCAAGTCCAT	4896
Qy	1641	LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro	1660
Db	4897	CTGAGNACTGCATCTGAAGATTTAAAGCCAGGTTCACAAAGTCAATCTGTTCTGTGATCCA	4956
Qy	1661	GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln	1680
Db	4957	GGCTTCCAGCTGGTGGGAACCTGTGTGAGTACTGTCTGAATCAGAGCAGACAGTGCACAA	5016
Qy	1681	ProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPhe	1700

Db	5017	CCACTTCTCTCACTGTGAACGCAATTAGCTGTGGGTGCCACCTCTCTTTTGGAGATGGCTTC	5076
Qy	1701	HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr	1720
Db	5077	CATTTCAGCCGATGACTTCTATGCTGGCAGCAGACAGTAACCTACCAGTGAACAATGGCTAC	5136
Qy	1721	TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer	1740
Db	5137	TATCTATTGGGTGACTCAAGGATGTTCTGTACAGATAAATGGGAGCTGGAAACGGCGTTCA	5196
Qy	1741	ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer	1760
Db	5197	CAATCTGCTGATGTGCGATGAGTGTGCGATGGATGATGATGATGATGATGATGATGATGAT	5256
Qy	1761	CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly	1780
Db	5257	TGCCTGAACGTAGATGATCTACATATGTTCAATGTTGCCATGCCAGTACACAGGAGATGG	5316
Qy	1781	LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer	1800
Db	5317	AAAAACGTGTCAGAACCTTATAAATGTAAGGCTCCAGGAAATCCGAAATGGCCACTCC	5376
Qy	1801	SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln	1820
Db	5377	TCAGTGAAGTTTATACAGTAGTGTCCGAGTCAATTTTGTGTGTCAGGAGGATACCAG	5436
Qy	1821	LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro	1840
Db	5437	TTGATGGGAGTAACCAAAATCACATGTTTGGAGTCTGGAGATGGAATCATCTAATACCA	5496
Qy	1841	TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu	1860
Db	5497	TATTTGTAAGCTGTTTTCATGTGTAACCGGCTTATTCAGAAATGGTTCATTTGAGGAG	5556
Qy	1861	LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla	1880
Db	5557	TTAGCATTTACTTTTGGCAGCAAGTGCATATAGTGTATTAAGGATATATCTCTGGCC	5616
Qy	1881	GlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCys	1900
Db	5617	GGTGAATAAGAAATCATCTCTGTCTTGTCTAACAGTCTTTTGGAGTCAATTCCTCTCTGTGT	5676
Qy	1901	GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly	1920
Db	5677	GAACAGTGAAGTGTCTAGTCCGAAATATAAATAATAGGAAATATATTTTGGTGGG	5736
Qy	1921	LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro	1940
Db	5737	CTTACTTACTTCTACTCATCATCATATTCATGCGATACAGGATACAGCTTACAGGGCCCT	5796
Qy	1941	SerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeu	1960
Db	5797	TCCATTTATTTGAATGCAAGGCTTCTGGCATCTGGGACAGAGCGCCACTGCTGTGCACCTC	5856
Qy	1961	ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr	1980
Db	5857	GTCCTTCTGTGGAGAACCTGCGCATCAAGATGCTGTCAATACGGGGAATAATCTTCACT	5916
Qy	1981	PheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr	2000
Db	5917	TTTCAGGAAACACCGTCACTTACATTCGAAAGGCTATATCTCTGTGCTGTGACACC	5976
Qy	2001	IleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSer	2020
Db	5977	ATTGAATGCTGGCCGACGCAAGTGGAGTAGAAGTACAGCAGTGTGCTGCTGTCTCC	6036
Qy	2021	CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly	2040
Db	6037	TGTGATGAGCCACCATTTGTGGACCAAGCTCTCCAGAGACTGCCCATCGGCTCTTTGA	6096
Qy	2041	AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu	2060
Db	6097	GACATTTGCTTCTACTTACTTCTGATGTTTACAGCTAGCAGCAATTCACAGCTTCTC	6156

2061 CysAsnAlaGlnGlyLysTrpValProProGluGlnAspMetProAsnCysIleala 2080
Db TGCATGCCAGGCGAAGTGGGTACCCCAAGAGTCAAGACATGCCCCGGTGTATAGCT 6216
2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerIleA 2100
Db CATTTCTGTGAAAACCTCCATCGGTTCCTATAGCATCTTGGAACTCTGTGAGCAAGCA 6276
2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
Db AAAATTGCAGCTGGCTCAGTTGTGAGCTTAAATGTCATGGAAGGCTTTGTACTGAACACC 6336
2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
Db TCACAAGATTTGAATGTATGAGAGGTGGCAGTGGAAACCTTCCCCCATGTCCATCCAG 6396
2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
Db TGCATCCCTGTGCGGTGTGGAGAGCCCAAGCATCATGATGGCTATGCAAGTGGATCA 6456
2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
Db AACTACAGTTTTCGAGGCATGGTGGCTTACAGCTGCAACAAAGGGTTCTACATCAAGGG 6516
2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerProIleProThrCysHis 2200
Db GAAAAGAGAGACACTCGCAAGCCACAGGCGAGTGGAGTAGTCTTATACCGACGTGCCAC 6576
2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
Db CCGGTATCTTGTGTGTAACCACTAAGTTGAGAAATGGCTTCTTGAGGCATACAACTGGC 6636
2221 ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer 2240
Db AGGATCTTTCAGAGTGAAGTGAAGTATCAGTGTAAACCGGGCTATAAGTCAAGTGGAGT 6696
2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
Db CCTGTATTTGTGTGCCAAGCCATCGCCATCGGCACAGTGAATCCCTCTCATGTGTGT 6756
2261 ProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe 2280
Db CCTCTCGACTGTGAAAACCTCCCGCATCCAGAAATGGCTTTCATGAAAGGAGAAACTTT 6816
2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
Db GAAAGTAGGGTCCAAAGTTCAGTTTCTGTAAATGAGGGTTATGAGCTTGTGTGACAGT 6876
2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320
Db TCTTGGACATGTCAAAATCTCGCAATGGAAATGAAGATCAATCCAAAGTGCATGCCCT 6936
2321 AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThr 2340
Db GCCAAGTGCCAGAGCCGCCCTCTTGAAAAACCCAGCTAGTATTAAAGGAGTTGACCCACC 6996
2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
Db GAGTAGGAGTTGTGACATTTCTCTGAAAGAGGCGATGTCTCTGCAAGGCCCTCTGTCTC 7056
2361 LeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
Db CTGAAATGCTGTCCTCCCAAGAAAGAAAGTACTCTTCTCTGTTGTGAAGATGTTCTT 7116
2381 CysThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
Db TGTACCCCACTCCCTCAATTTCTTTGGTGTCCCAATCTCTCTCTCTCTCTCTCTCTCT 7176
2401 GlySerThrValLysTyrSerCysValGlyGlyPheLeuArgGlyAsnSerThrThr 2420
Db GGAAGTACTGTCAAGTATCTTGTGTAGTGGGTTTTTCTTAAGAGGAAATTTCTACCAACC 7236
2421 LeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys 2440
Db CTCCTGCCAACTGATGGACCTCTGGAGCTCTCCACTGCCAAGATGTGTTCCAGTAGATGT 7296
2441 ProGlnProGluGluIleProAsnGlyIleLeuAspValGlnGlyLeuAlaTyrLeuSer 2460
Db CCCCAACCTGAGGAATCCCAATGGAATCATTTGATGTGCAAGGCTTGCTCTATCTCAGC 7356
2461 ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys 2480
Db ACAGCTCTCTATACCTGCAACGAGGCTTTGAATTTGGTGGGAATACTACCAACCTTTGT 7416
2481 GlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLys 2500
Db GGAGAAATGTCTACTGCTTGGAGGAAACCAACATGTAAGCCATTGAGTGCCTGAAA 7476
2501 ProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal 2520
Db CCCAAGGAGATTTTGAATGGCAATTTCTTACACGGACCTACACTATGGACAGACCGTT 7536
2521 ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu 2540
Db ACCTACTCTTCCAAACCGAGGCTTTGGCTCGAAGGTCCAGTCCCTTGACCTGTTTAGAG 7596
2541 ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln 2560
Db ACAGTGTATGGAGTAGATGCCCATCTTGCATGTCATCCACTGTGATTTCCCAACA 7656
2561 ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleLys 2580
Db CCCATGAAAATGTTTTGTAGAAGTGCAGATTACAGTATGTTGCCATATCATCTACTAC 7716
2581 SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluLysSerGly 2600
Db AGTTGCTTCCCTGGGTTTCAGTGGCTGGTCAATGCATGAGACTGTGGCTTCCCTCATATAGAT 7776
2601 TrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp 2620
Db TGGTCAAGTTCCATCCCAACATGATATGCCATAGACTGTGGCTTCCCTCATATAGAT 7836
2621 PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMet 2640
Db TTTTGAGACATGCTACTAACTCAAAGATGACCCAGGATATTTTGAGCAAGAGACGACATG 7896
2641 MetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThr 2660
Db ATGGAAGTTCCATATGAGCTCTCCCTTATCAATTTGGAGAGAGTGGCTAAAAACC 7956
2661 TrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet 2680
Db TGGGAATACAAAGGAGTCTCTCTGTGTACACATTCATCAAACTTTCTGTATGTTACCATG 8016
2681 ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGln 2700
Db GTTTTCATACACTCTGTAATCCAGGATATGAATCTCTGGGGAAACCCCTGTGTCTATCTGCCAG 8076
2701 GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
Db GAAATGGAATCTGGAATGGCGATGACCATCTCTGCAATTTCAATTTGAATGTGACTTGCT 8136
2721 ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln 2740
Db ACTGCTCTGAAATGGCTTTTGGTTTTTACAGAGACTAGCATGGGAAGTGTCTGTGCAG 8196
2741 TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
Db TATAGCTGTAAACCTGGACACATTTCTAGCAGGCTCTGACTTTAAGGCTTTTGTCTAGAGAT 8256
2761 ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
Db AGAAGTGGAGTGTGCTCCCGCTGTAAGCCATTTTATGCAAAAGCAATTCOA 8316
2781 ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu 2800

2061 CysAsnAlaGlnGlyLysTrpValProProGluGlnAspMetProAsnCysIleala 2080
Db TGCATGCCAGGCGAAGTGGGTACCCCAAGAGTCAAGACATGCCCCGGTGTATAGCT 6216
2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerIleA 2100
Db CATTTCTGTGAAAACCTCCATCGGTTCCTATAGCATCTTGGAACTCTGTGAGCAAGCA 6276
2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
Db AAAATTGCAGCTGGCTCAGTTGTGAGCTTAAATGTCATGGAAGGCTTTGTACTGAACACC 6336
2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
Db TCACAAGATTTGAATGTATGAGAGGTGGCAGTGGAAACCTTCCCCCATGTCCATCCAG 6396
2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
Db TGCATCCCTGTGCGGTGTGGAGAGCCCAAGCATCATGATGGCTATGCAAGTGGATCA 6456
2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
Db AACTACAGTTTTCGAGGCATGGTGGCTTACAGCTGCAACAAAGGGTTCTACATCAAGGG 6516
2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerProIleProThrCysHis 2200
Db GAAAAGAGAGACACTCGCAAGCCACAGGCGAGTGGAGTAGTCTTATACCGACGTGCCAC 6576
2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
Db CCGGTATCTTGTGTGTAACCACTAAGTTGAGAAATGGCTTCTTGAGGCATACAACTGGC 6636
2221 ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer 2240
Db AGGATCTTTCAGAGTGAAGTGAAGTATCAGTGTAAACCGGGCTATAAGTCAAGTGGAGT 6696
2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
Db CCTGTATTTGTGTGCCAAGCCATCGCCATCGGCACAGTGAATCCCTCTCATGTGTGT 6756
2261 ProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe 2280
Db CCTCTCGACTGTGAAAACCTCCCGCATCCAGAAATGGCTTTCATGAAAGGAGAAACTTT 6816
2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
Db GAAAGTAGGGTCCAAAGTTCAGTTTCTGTAAATGAGGGTTATGAGCTTGTGTGACAGT 6876
2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320
Db TCTTGGACATGTCAAAATCTCGCAATGGAAATGAAGATCAATCCAAAGTGCATGCCCT 6936
2321 AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThr 2340
Db GCCAAGTGCCAGAGCCGCCCTCTTGAAAAACCCAGCTAGTATTAAAGGAGTTGACCCACC 6996
2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
Db GAGTAGGAGTTGTGACATTTCTCTGAAAGAGGCGATGTCTCTGCAAGGCCCTCTGTCTC 7056
2361 LeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
Db CTGAAATGCTGTCCTCCCAAGAAAGAAAGTACTCTTCTCTGTTGTGAAGATGTTCTT 7116
2381 CysThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
Db TGTACCCCACTCCCTCAATTTCTTTGGTGTCCCAATCTCTCTCTCTCTCTCTCTCTCT 7176
2401 GlySerThrValLysTyrSerCysValGlyGlyPheLeuArgGlyAsnSerThrThr 2420
Db GGAAGTACTGTCAAGTATCTTGTGTAGTGGGTTTTTCTTAAGAGGAAATTTCTACCAACC 7236

DB 8317 GTTCATGATGGATCCATCAAGGAGCACTACATACCTGAGCAGCTGTACTATGAG 8376
QY CysAspProGlyThrValLeuAenGlyThrGluArgThrCysGlnAspAspLysAen 2820
DB 8377 TGTGACCCCGGATATGTGCTGAATGCATCTGAGAGGAGAACATGCGCAGGATGACAAAAAC 8436
QY TTPAspGluAspGluProIleCysIleProValAspCysSerProProValSerAla 2840
DB 8437 TGGGATGAGGATGAGCCCATTTGCTCTGTGAGCTGAGTTTCAACCCCGCTCAGCC 8496
QY AenGlyGlnValArgGlyAspGluThrPheGlnLysGluIleGluThrCysAen 2860
DB 8497 AATGCCAGGTGAGGAGAGCGATACATTTCCAAAAGAGATTGAATACACTTGCAT 8556
QY GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAenGlySerTrpSer 2880
DB 8557 GAAGGGTTCTTGCTTGAGGAGCCAGGAGCTCGGTTTGTCTTCCCAATGGAAGTTGGAGT 8616
QY GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAenGly 2900
DB 8617 GGAGCCACTCCCGACTGTGTGCTGTGATGTGCCACCCCGCCCAACTGGCCAAATGGG 8676
QY ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly 2920
DB 8677 GTGACGGAAGGCTGGACTATGGCTTCATGAGGAAGTAACTTCACCTGTACAGAGGC 8736
QY TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu 2940
DB 8737 TACATCTTGCACGGTCTCCAAAACCTCACCTGTGCTCAGTACAGTGGCACTGGGATGCAGAG 8796
QY IleProLeuCysLysProValAenCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
DB 8797 ATTCTCTCTGTAAACCGTCAACTGTGACCTCTCGAAGATCTTCCCATGTGTTCCCT 8856
QY AenGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980
DB 8857 AATGGTTTTCCTTTATTCATGCGGCGCCATATACAGTATCAGTGTCTTCTGGTTATAG 8916
QY LeuHisGlyAenSerSerArgArgCysLeuSerAenGlySerTrpSerGlySerPro 3000
DB 8917 CTCCATGGAATTCATCAAGAGTGCCTCTCCAAATGGCTCTCCAGAGTGGCAGCTCACT 8976
QY SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAenGlyThr 3020
DB 8977 TCTGCTGCTGCTGCGAGTGTCCACACAGTAAATGAATATGGAATCTGTCAATGGGACA 9036
QY AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
DB 9037 GATTTTGACTGTGGAAGGCGAGCCCGATTCAGTGTCTCAAGGCTTCAAGCTCTTAGGA 9096
QY LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu 3060
DB 9097 CTTTCTGAAATCACCTGTGAAGCGGATGCGCAGTGGAGCTCTGGATTCCCCCACTGTGAA 9156
QY HisThrSerCysGlySerLeuProMetIleProAenAlaPheIleSerGlnThrSerSer 3080
DB 9157 CACACTTCTGTGGTTCTCTTCCAAATGATACCAATGCGTTCAATCAGTGAACAGCTCT 9216
QY TrpLysGluAenValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
DB 9217 TGGAGGAAATGTGATTAATTAACACTGACGTGCGGTCTGGATATGTATACAGGCACTTCA 9276
QY AspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeu 3120
DB 9277 GATCTGATTTGTACAGAGAAAGGGTATGAGGCGAGCTTATCCAGTCTGTGAGCCCTTG 9336
QY SerCysGlySerProProSerValAlaAenAlaValAlaThrGlyGluAlaHisThrTyr 3140
DB 9337 TCTGTGGGTCCCCACCGTGTGCGCAATGCAAGTGGCACTGAGGAGGCACACCTAT 9396
QY GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThr 3160

DB 9397 GAAAGTGAAGTGAACCTCAGATGCTCGAAAGGTTATACGATGATACAGATACAGATACA 9456
QY PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys 3180
DB 9457 TTCCACCTGTGCAGAAAGATGTGCTGCTGCTCCCTGAGAGAAATCTCTGCGAGTCTTAAAAA 9516
QY CysProLeuProGluAenIleThrHisIleLeuValHisGlyAspAspPheSerValAen 3200
DB 9517 TGTCTCTCTCCCGGAAAAATACATACATATCTTGTATCATGGGACGATTTCACTGTGAAT 9576
QY ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAenIleSerVal 3220
DB 9577 AGGCAAGTTTCTGTGTCTCATGTGCAGAAAGCTATACCTTTTGGAGGAGTTAAATACATACAT 9636
QY CysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSer 3240
DB 9637 TGTGAGCTGTGATGGAACCTGGGAGGCCCATTTCTCGATGAATCTTTGCGAGTCCAGTTCT 9696
QY CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
DB 9697 TGTGGGAAACCTGAAAGTCCAGAAACATGATTTGTGTGGCAGTAAATACATACCTTTGAA 9756
QY SerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAenArgGluArgVal 3280
DB 9757 AGCACAAATTTATTCAGTGTGAGCTGCTGCTATGAACTAGAGGGCAACAGGAAACGCTGC 9816
QY CysGlnGluAenArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGlu 3300
DB 9817 TGCAGGAGAACAGACAGTGGAGTGGAGGGTGGCAATATGCAAGAGAGACAGGTGTGAA 9876
QY ThrProLeuGluPheLeuAenGlyLysAlaAspIleGluAenArgThrThrGlyProAen 3320
DB 9877 ACTCCACTTGAATTTCTCAATGGGAAAGCTGACATTTGAAACAGGAGCACTGGACCCAC 9936
QY ValValTyrSerCysAenArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
DB 9937 GTGGTATATTCCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCAACCTGCACA 9996
QY GluAenGlyThrTrpSerHisProValProLeuCysLysProAenProCysProValPro 3360
DB 9997 GAAATGGAACCTGGAGGCCACCCAGTCCCTCTCTGCAAAACCAATCCATGCCCTGTTCCT 10056
QY PheValIleProGluAenAlaLeuLeuSerGluLysGluPheTyrValAspGlnAenVal 3380
DB 10057 TTTGTGATTTCCCGAGAGATGCTCTGCTGTCTGAAAGAGGATTTTATGTTGATCAGATGTG 10116
QY SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAenPro 3400
DB 10117 TCCATCAATGTAGGGAAGGTTTCTGCTGCGAGGCCACGCGCATCATTTACCTGCAACCCC 10176
QY AspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAla 3420
DB 10177 GACGAGACGTGGACACAGACAGCGCCCAATGTGAAAAAATCTCATGTGTGTCCACAGCT 10236
QY HisValGluAenAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
DB 10237 CAGTAGAATAATCAATTTCTCGAGGGGTACATTTATCAATATGAGACATGATCACCTAC 10296
QY SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAenGly 3460
DB 10297 TCATGTTACATGCTGATACATGTTGGAGGGTTCTCTGAGGAGTGTCTGTTTGTAGAAAATGGA 10356
QY ThrTrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAenGlyGly 3480
DB 10357 ACATGAGCATCACCTCTTATTTGACAGCTGTCTGTGATTTCCATGTGAGATGGGGGC 10416
QY IleCysGlnArgProAenAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGlu 3500
DB 10417 ATCTGCAACGCCCAATGCTTGTCTGTCCAGAGGGCTGGATGGGGCCCTCTGTGAA 10476
QY GluProIleCysIleLeuProCysLeuAenGlyGlyArgCysValAlaProTyrGlnCys 3520
DB 10477 GAACCAATCTGCATTTCTCTCTGTCGAACGGAGGTGCTGTGTGTGGGCCCTTACAGTGT 10536

QY 3521 AspCysProGlyTrrpThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
Db 10537 GACTGCCCGCTGCTGGAGGGGCTCTCGCTGTCTATACAGCTGTTCAGCTCTCCCTGC 10596
QY 3541 LeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThrGly 3560
Db 10597 TTAATGGTGGAAATGTGTAAAGCAACCGATGTCTCTTCTTCTTGGACGGGA 10656
QY 3561 HisAsnCysSerArg 3565
Db 10657 CATAACTGTTCACGG 10671

RESULT 5

US-10-028-248A-7
; Sequence 7, Application US/10028248A
; Publication No. US20030235882A1
; GENERAL INFORMATION:
; APPLICANT: Shimketa, Richard
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine
; APPLICANT: Casman, Stacie
; APPLICANT: Malyankar, Uriel
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gangoli, Esha
; APPLICANT: Miller, Charles
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Smithson, Glennda
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Liu, Xiaohong
; APPLICANT: Colman, Steven
; APPLICANT: Tchernev, Velizar
; APPLICANT: Si, Jingsheng
; APPLICANT: Edinger, Shlomit
; APPLICANT: Stone, David
; APPLICANT: Sciore, Paul
; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods
; TITLE OF INVENTION: Thereof
; FILE REFERENCES: 21402-222
; CURRENT APPLICATION NUMBER: US/10/028,248A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 11152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-028-248A-7

Alignment Scores:

Pred. No.:	0	Length:	11152
Score:	19616.00	Matches:	3510
Percent Similarity:	98.82%	Conservative:	16
Best Local Similarity:	98.37%	Mismatches:	40
Query Match:	98.21%	Indels:	2

16 Gaps: 1

DB: US-09-977-053-4 (1-3571) x US-10-028-248A-7 (1-11152)

QY 4 ArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThrPheGlnGln 23
Db 83 AGAATTTGCGCGGCTGCTGGGCTCTGGCGCTCGTTTGGGCTGGCGACCTTTTCAGCAG 142
QY 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db 143 ATGTCCCGCTGGCGCAATTTTCAGCTTCCGCTCTTCCCGAGACCGCGCGCGGCCCCC 202
QY 44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
Db 203 GGGAGTATCCCGCGCGCGCGCTCTTCCGACGACGAGCGCGGGGAGCAGAGTGGAGCGG 262
QY 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
Db 263 CTGGGCCAGCGCTTC-----CGCGTGGCTGCTGGCGGAGCTCAGCGAGCGCTGGAG 316
QY 84 LeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSerGluLeu 103
Db 317 CTGTCTTCTCTGGTGGAGATTCGTCCAGCGTGGCGGAAGTCAACTTCGCGAGCGAGCTC 376
QY 104 MetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db 377 ATGTTCTGTCGCAAGCTGCTGTCCGACTTCCCGTGGTGGCCACCGCCACCGCGCGTGGCC 436
QY 124 IleValThrPheSerSerLysAsnTrpValValProArgValAspTrpIleSerThrArg 143
Db 437 ATCTGTACCTTCTCGTCCAGAACTACGTGTGGTGGCGCGCTCGATTACATCTCCACCGCG 496
QY 144 ArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluLeuProAlaIleSerTrpArg 163
Db 497 CGCGCGCGCCAGCAAGTGGCGCTGCTCTCTCAAGAGATCCCTGCCATCTCTTACCGA 556
QY 164 GlyGlyGlyThrTrpLysGlyValaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
Db 557 GGTGGCGGCACCTACACACAGGGCGCTTCCAGCAAGCGCGCAAAATCTCTTCTCATGCT 616
QY 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTrpSerAsnGlyGlyAsp 203
Db 617 AGAGAAATCTCAACAAAATTTGTATTTCTCATCTACTGATGGATATTCCAAATGGGGAGAC 676
QY 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db 677 CCTAGACCAATTTGACGGCTCACTCGAGATTCAGAGTGGAGATCTTCACTTTTGGCATA 736
QY 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
Db 737 TGGCAAGGGACATTCGAGAGCTCAATGACATGGCTTCCACCCCAAGGAGGAGCACTGT 796
QY 244 TyrLeuLeuHisSerPheGluPheGluAlaLeuAlaArgArgAlaLeuHisGluAsp 263
Db 797 TACCTGCTACACAGATTTTGAAGAAATTTAGCTCGCGGGCATTTGCATGAAGAT 856
QY 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrpLeuCysAspGlu 283
Db 857 CTACCTTCTGGGAGTTTATTCAGAGATGATGTGCTCACTGCTCATATCTTTGTGATGAG 916
QY 284 GlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303
Db 917 GGCAGGACCTGCTGTGACCGCAATGGGAAGCTGCAAAATGTGGGACACACACAGGCCATTTT 976
QY 304 GluCysIleCysGluLysGlyTyrGlyLysGlyLeuGlnTrpGluCysThrAlaCys 323
Db 977 GAGTGCATCTGTGAAAAGGGGTATTACGGGAAAGGTCTGCAAGTATGAATGCACAGCTTGC 1036
QY 324 ProSerGlyThrTrpLysProGluGlySerProGlyGlyIleSerSerCysIleProCys 343
Db 1037 CCATCGGGGACATCAAAACCTGGAAGCTTCAACAGAGGAGGAGGAGGAGGAGGAGGAGG 1096
QY 344 ProAspGluAsnHisThrSerProGlySerThrSerProGlySerProGluAspCysValCysArg 363

1097 CCGGATGAATATCACACCTCTCCACCTGGAGACACATCCCTCGAGACTGTGTCTGAG 1156
364 GluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
1157 GAGGATACAGGCGATCTGCCAGACCTGTGAACCTGTGTCACCTGCCCTGGCCCTGAAGCCT 1216
384 ProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGly 403
1217 CCGGAAATGGTTACTTTATCCAAACACTTGCACACACACTTCATGAGCGCTGTGGG 1276
404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsn 423
1277 GTCCGATGTACCCCTGGATTTGATCTTGTGGAAGCAGCATCATCTTATGTCTACCCAAT 1336
424 GlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGln 443
1337 GGTTTGTGGTCCGGTTTCAGAGAGTACTGCAGAGTAAGACATGTCTCTCATCTCCGCCAG 1396
444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeu 463
1397 CCGAAACATGCCACATCAGCTGTCTTCTACAGGGAATGTTATATAGACACATGTTTG 1456
464 ValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGlyAsn 483
1457 GTTGCTGTGATGAGGGTACAGACTAGAGGCAGTGTATAGCTTACTTGTCAAGGAAAC 1516
484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
1517 AGCCAGTGGAGTGGCGCAGAACCCCGGTGTGTGGAGCGCCACTGTCTTCACATG 1576
504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
1577 CCCAAAGATGTCTATCATATATCCCCCACTCTGTGGCAGCAGCAGCAAAATTTGGAGCG 1636
524 IleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
1637 ATCTGCTATGTAAGTTGCCGCCAAGGGTTCAATTTATCTGGAGTCAAGAAATGCTGAGA 1696
544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLysAspValGlu 563
1697 TGTACCACCTTCGGAAATGGAATGTGGAGTTCAGGCAGCTGTGTGTAAAGACGTGGAG 1756
564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlnGlnAspSer 583
1757 GCTCCTCAAAATCAACTGTCTCAGGACATAGAGCTTAAGACTCTGGAAACAGCAAGATTCT 1816
584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerVal 603
1817 GCCAATGTTACCTGGCAGATTCACACAGCTAAAGACAACTCTGTGTGAAAGGTGTCAATC 1876
604 HisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleVal 623
1877 CGCGTTTCATCCAGCTTTTACCCCACTTACCTTTTCCCAATTTGGAGATGTGTATCGTA 1936
624 TyrThrAlaThrAsnLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
1937 TACCGGCAACTGACCTATCCGCAACACAGGCGAGCTGCATTTTCATATCAAGGTTAT 1996
644 AspAlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGlu 663
1997 GATGCAAGACCACTGTCTAGACTGTGTGAGATCTCCACTCCCGTCCAGGTTCTGGAG 2056
664 LysValHisAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu 683
2057 AAGGTACATGCGCAAGCTGGATGAGCTCTCAGTTCTCAGACAACTCAGGGGCTGAATTG 2116
684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
2117 GTCAITACCAAGAGTATACACAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAG 2176
704 TyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLys 723

2177 TATACAGCCACTGACCCCTCAGGTAATAACAGGATATGTGTATATTCATATGTGTCAATAA 2236
724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
2237 GGTTCCTCCCTGTGAATTCATATTCACACCTGTAAATGGGATTTTATATGCACCTCCAGAT 2296
744 AsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
2297 AATCTGGAGTCAACTGTACATTAACCTTGTGGAGGGCTACGATTTTCACAGAGGGTCT 2356
764 ThrAspLysTyrTyrCysAlaLysGluAspGlyValTrpLysProThrTyrThrGlu 783
2357 ACTGCAAGATTAATTTGTCTTATAGAGATGGCGCTCTGGAACCCACATATACCACTGAA 2416
784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
2417 TGCCAGACTGTGCCAAAACGTTTTTGCAAAACACCGGGTTCAAGTCTCTTTGAGATGTT 2476
804 TyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu 823
2477 TACAAGCAGCTGTTGTGATGACTCAGATCTGATGAAGAAGTTTTCTGAAGCATTTGAG 2536
824 ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
2537 ACAGCCCTGGCAAAAATGGTCCCATCATTTTGTAGTCAATGCAGAGCAATTCAGCTGAGA 2596
844 LeuGluLeuAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly 863
2597 CTGGAGAGAACCTGACCAAAAATATTTGCTAGATATATTAATATGACTATGAANAATGGC 2656
864 PheAlaIleGlyProGlyGlyTyrGlyValAlaAsnArgLeuAspTyrSerTyrAspAsp 883
2657 TTTGCAATTTGTCAGGTGGCTGGGTGAGCTAATAGGCTGGATTACTCTTACGATGAC 2716
884 PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle 903
2717 TTCTTGACACTGTGCAAGAAACAGCCACCAAGCATCGCAATGCGCAATGCTCCAGGATT 2776
904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
2777 AAAAGAAGTGGCCCATTTATCTGACTATAAATTAAGTTAATTTTAAACATCACAGTAGT 2836
924 ValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeu 943
2837 GTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAATTCAGCAACGACTCCTT 2896
944 GlnThrLeuGluThrIleThrAsnLysLeuLysLeuArgThrLeuAsnLysAspProMetTyr 963
2897 CAGACATTTGGAACCTATCAAAAATAAATCTGAAAAGGACTCTCAACAAAGACCCCATGTAT 2956
964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAsnSerAsnSerLeuGluThrLysLys 983
2957 TCTTTTCAGCTTGCATCAGAAATACTTTATAGCCAGCAGCAATTCATTAGAAACAAANAAG 3016
984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003
3017 GCTTCCCTCTCTGCAGACGAGCTCAGTCTGAGAGGGGTATGTGTGTCAATGGCCT 3076
1004 LeuGlyThrTyrThrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr 1023
3077 TTGGGAACCTTATTAATCTGGAACATTTTCACTGTGAAAAGCTGCCCGGATCGGATCTTAT 3136
1024 GlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyr 1043
3137 CAAGATGAAGAGGGCACTTGAGTGCAGCTTTTGGCCCTCTGGGATGTACCGGAATAT 3196
1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
3197 ATCCATTCAAGAAACATCTCTGATTTGAAGCTCAGTGTAAACAAGGCACCTCTACTATAC 3256
1064 SerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySer 1083
3257 AGTGGACTTGAGACTTTGTGAATCTGTGTCCACTGGGCACTTATTCAGCCAAAATTTGGTTCC 3316

QY	1084	ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle	1103
DB	3317	CGGAGCTGCCTCTCGTGTCCAGAAAAACACCTCAACTGTGTGAAAAGAGAGCGGTGAACATT	3376
QY	1104	SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys	1123
DB	3377	TCTGCATGTGGAGTTCTCTGTCCAGAAAGGAAAAATCTCTGGGGTTCTGGGGTTAAATGCCCTGT	3436
QY	1124	HisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys	1143
DB	3437	CACCCATGTCCTGTGTGACTATTATCAACCTAAATGCAGGAAGGCGCTTCTGCCCTGGCGCTGT	3496
QY	1144	ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPhe	1163
DB	3497	CCCTTTTATGGAACTACCCCAATTCGCTGGTTCAGATCCATCAGCAATGTTCAAGTTT	3556
QY	1164	SerSerThrPheSerAlaAlaGluGluSerValValProProlaSerLeuGlyHisIle	1183
DB	3557	AGTTCAACTTTCTCAGCGGCAGAGGAAGTGTGGTGGCCCCCTGCCTCTCTTGACATATT	3616
QY	1184	LysIlyAsnHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHis	1203
DB	3617	AAAAAGAGGCGATGAATCAGCAGTCAGCGCAAGTCATGAATGCTTCTTTAAACCTTGCCAC	3676
QY	1204	AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr	1223
DB	3677	ATAAGTGGAACTGCCAGCACTTGGCGGTGGTTATGTTTGTCTCTGTCCACTTGATAT	3736
QY	1224	ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsn	1243
DB	3737	ACAGGTTTAAAGTGTGAAACAGACATCGATGAGTGCAGCCCACTGCCTTGCCCTCAACAAT	3796
QY	1244	GlyValCysLysAspLeuValGlyLysPheIleCysGluCysProSerGlyTyrThrGly	1263
DB	3797	GGAGTTTGTAAAGACCTTAGTTGGGGAAATCATTTGTGTAGTGCCTCAGGTTACAGGT	3856
QY	1264	GlnArgCysGluGluLeuAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIle	1283
DB	3857	AAGCAGTGTGAATGAACATCAATGAATGTCAGTCTAATCCATGTAGAAATCAGGCCACC	3916
QY	1284	CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis	1303
DB	3917	TGTGTGGATGAATTAATTCATACAGTTGTAAATGTCCAGCCAGGATTTTCAGGCCAAAAGG	3976
QY	1304	CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGlu	1323
DB	3977	TGTGAACACAGGTATGTATCACTCAGTGTGTTTAAATAAACCTTAAATAGTCAGTCTGTGA	4036
QY	1324	AspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGly	1343
DB	4037	GACCAGGTTGGGGATTCCTGTGTGAAATGCCACCTGGATTTTGGGTACCCCATGTGGA	4096
QY	1344	LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly	1363
DB	4097	AAGNACTCGATGAGTGTCTCAGTCAGCCATGCBAATGAGACTACTCTAAGACGGT	4156
QY	1364	AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn	1383
DB	4157	GCCAAATAGCTTCAGTGCCTGTGTGCAGCTGGCTTCCAGGATCCACTGTGAATGAAC	4216
QY	1384	IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn	1403
DB	4217	ATCAATGAATGTCACTTAATCCCATGTAGAAATCAGGCCACCTGTGTGGATGAATTAAT	4276
QY	1404	SerTyrSerCysLysCysGlnProProGlyPheSerGlyGlnArgCysGluThrGluInSer	1423
DB	4277	TCATACAGTTGTAAATGTTCAGCCAGGATTTTCAGCCAAAAGGTGTGAACAGAACAGTCT	4336
QY	1424	ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly	1443
DB	4337	ACAGCCTTTAAACCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCATGCTAGATGGC	4396

Qy	1444	MetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetIysSerSerAspMet	1466
Db	4397	ATGCTCCCATCTCTCCATGCTCTAAACCTGCTACCTTCTGGATGAAATCTCTGTACGACATG	4456
Qy	1464	AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu	1483
Db	4457	AACATATGGAAACCAACATCTCTATGATGAGTGTAAACGGCAGCGACAACTACCTTGCCTCCTG	4516
Qy	1484	ThrAspTyrAsnGlyTyrValLeuLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro	1503
Db	4517	ACTGATTTAAACGGGTGGGTCTTTATGTGAAATGGCAGGGGAAAGATAACAACTGTCTCC	4576
Qy	1504	SerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAlaAsnGlyIle	1523
Db	4577	TCGGTGAATGATGGCAGATGGCATCATATTGCCAATCACTTGACAAAGTACTTGGTGAGGCC	4636
Qy	1524	TrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu	1543
Db	4637	TGGAGGCTCATATAAATGGGGAATATCTGACGGTGGTACTTGGCTCTCTCCATTGGCCAA	4696
Qy	1544	ProIleProGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysGlyGluGly	1563
Db	4697	GCCATACCTGGTGGCGGTGCATTAGTCTTCTGGCAGACAGACAAAGAGAGAGGGG	4756
Qy	1564	PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal	1583
Db	4757	TTCAACCGGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGACTATGTC	4816
Qy	1584	LeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGly	1603
Db	4817	CTGTCTCCACAGCAGGTGAAGTCACTGGCTACTCTCTGCCAGAGAGAACTCAAGTAAAGA	4876
Qy	1604	AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLysIleAspSer	1623
Db	4877	AACGTGTAGCATGGCGCTGATTTCTTGTGAGGAATGTGGGGAAGTGAAGATCGATTCT	4936
Qy	1624	LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr	1643
Db	4937	AACAGCATATTTTGTCTGATTTGCCACGCTTGGGAGGCTCAGTGCTCATCTGAGAACT	4996
Qy	1644	AlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln	1663
Db	4997	GCATCTGAAGAGATTAAACCCAGGTTCCAAAGTCAATCTGTTCTGTGAACCAAGCTTCCAG	5056
Qy	1664	LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuPro	1683
Db	5057	CTGGTCGGGAACCTGTGCAGTACTGTCTGAATCAGGACAGTGGACACACCACTCTCCC	5116
Qy	1684	HisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHisSerAla	1703
Db	5117	CACGTGTGAACGCAATTCGCTGTGGGTGCCACCTCTCTTTGGAGAAATGGCTTCCATTGAGCC	5176
Qy	1704	AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu	1723
Db	5177	GATGACTTCTATGCTGGCAGCACAGTAACCTACCAAGTGCACAAATGGCTACTATCTATTG	5236
Qy	1724	GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys	1743
Db	5237	GGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCGTTTACCATTCTGCG	5296
Qy	1744	LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn	1763
Db	5297	TTAGATGTCGATCAGTGTGCAGTTGGATCAGATTGTAGTGAGCATGCTCTTCTGGCTGAAC	5356
Qy	1764	ValAspGlySerTyrIleCysSerCysValProProTyrTrpThrGlyAspGlyLysAsnCys	1783
Db	5357	GTAGATGGATCTTACATATGTTTCATGTGTCCACCGTACACAGAGATGGGAAAACTGT	5416
Qy	1784	AlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu	1803
Db	5417	GCAGAACCTATAAAATGTAAAGGCTCAGGAAATCCCGGAAATGGCCACTCTCTCAGGTGAG	5476
Qy	1804	IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly	1823

Db 5477 ATTTATACAGTAGGTCGGAAGTACATTTTCGGTTCAGGAAGATACAGTTGATGGGA 5536
Qy 1824 ValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyrCysLys 1843
Db 5537 GTAACCAAAATCACATGTTTGGAGTCTGGAGATGGAATCATCTAATACCATATTGTAA 5596
Qy 1844 AlaValSerCysGlyProAlaIleProGluAsnGlyCysIleGluGluLeuAlaPhe 1863
Db 5597 GCTGTTTCATGTGTAAACCGGCTATTCCAGAAATGGTTGCATTGGAGAGTTAGCATTT 5656
Qy 1864 ThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLys 1883
Db 5657 ACTTTTGGCACAAGTGCATATAGGTGTAATAAGGATATCTCTGGCCGGTGATAAA 5716
Qy 1884 GluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProValCysGluProVal 1903
Db 5717 GAATCATCTGCTGTCTTAACAGTTCTTGGAGTCATTCCTCCCTGCTGTGTGAAACAGTG 5776
Qy 1904 LysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeuThrTyr 1923
Db 5777 AAGTGTTCCTAGTCCGGAAATATAAATAATGGAATAATATATTTTGAGTGGGCTTACCTAC 5836
Qy 1924 LeuSerThrAlaSerTyrCysAspThrGlyTyrSerLeuGlnGlyProSerIleIle 1943
Db 5837 CTTTCTACTGCATCATATTCATGCGTACAGATACAGCTTACAGGGCCCTTCCATTAAT 5896
Qy 1944 GluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCys 1963
Db 5897 GAATGCACGGCTTCTGGCATCTGGGACAGAGCGCCACCTGCTGTCACCTGCTCTTCGT 5956
Qy 1964 GlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsn 1983
Db 5957 GGAAACACCATGCTCCATCAAGATGCTGTCTATTACGGGGAATAACTTCACCTTTCAGGAAC 6016
Qy 1984 ThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCys 2003
Db 6017 ACCGTCACTTACACTTGCAGAAAGGCTATATCTCTGCTGGTCTTGACACCATTTGATGC 6076
Qy 2004 LeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCysAspGlu 2023
Db 6077 CTGGCCGACGCGCAAGTGGAGTAGAAGTCAACGACAGTGCCTGGCTGTCTCTGTGTAGAG 6136
Qy 2024 ProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAspIleAla 2043
Db 6137 CCACCCATTTGTGACACACGCTCTCCAGAGATGCGCCATCGGCTCTTTGGAGACATTCGA 6196
Qy 2044 PheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAla 2063
Db 6197 TTCCTACTGCTCTGTATGGTTACGCTTACGAGCAATTCACGATTCCTGTGCAATGCG 6256
Qy 2064 GlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAlaHisPheCys 2083
Db 6257 CAGGCAAGTGGGTACCCCAAGAGTCAAGACATGCCCCCGTGTGTATAGCTCATTTCTGT 6316
Qy 2084 GluLysProProSerValSerTrpSerIleLeuGluSerValSerLysAlaIlePheAla 2103
Db 6317 GAAAAACCTCCATCGGTTTCTATAGCATCTTGGAAATCTGTGAGCAAGCAAAATTTGCA 6376
Qy 2104 AlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLys 2123
Db 6377 GCTGGCTCAGTTGTGAGCTTTAAATGCATGAGAGGCTTTGTACTGACACACCTCAGCAAG 6436
Qy 2124 IleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGlnCysIlePro 2143
Db 6437 ATTGAATGTATGAGAGTGGGCGAGTGAACCTTCCCTCCATGCTCCATCCAGTGCATCCCT 6496
Qy 2144 ValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSer 2163
Db 6497 GTGCGGTGTGAGAGCCCAAGCATCATGATGGCTATGCAAGTGGATCAAACTACAGT 6556
Qy 2164 PheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyLysLys 2183

Db 6557 TTTGGAGCCATGGTGGCTTACAGCTGCAACAGGGGTTTCTACATCAAGGGCGAAAGAAG 6616
Qy 2184 SerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisProValSer 2203
Db 6617 AGCACCTCGGAAGCCACAGGGCAGTGGAGTAGTCTCTATACCGACGTGCCACCGGTATCT 6676
Qy 2204 CysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePhe 2223
Db 6677 TGTGTGAAACCACTTACAGTTGAGATGGCTTTCTGGAGCATACAACCTGGCAGGATCTTT 6736
Qy 2224 GluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPhe 2243
Db 6737 CAGAGTGAAGTGAGCTATCAGTGTAAACCGGCTTATAAGTCAGTCGGAAGTCTCTGTATTT 6796
Qy 2244 ValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValProLeuAsp 2263
Db 6797 GTCTGCCAAGCAATTCGCCACTGGGCACAGTGAATCCCTCTGATGTGTGTTCTCTCGAC 6856
Qy 2264 CysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGly 2283
Db 6857 TGTGGAAACCTTCCCCGATCCAGATGGCTTCATGAAAGGAGAAACCTTTGAAGTAGGG 6916
Qy 2284 SerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTrpThr 2303
Db 6917 TCCAAGGTTCAAGTTTCTGTAAATCAGGCTTATGAGCTTGTGTGTCACAGTCTCTGGACA 6976
Qy 2304 CysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAlaLysCys 2323
Db 6977 TGTTCAGAAATTCGGCAATGGAATGAAGTCAAAATCAAAAGTGCATGCTGCCAAGTGC 7036
Qy 2324 ProGluProProLeuLeuGluAsnGlnLeuValLysGluLeuThrThrGluValGly 2343
Db 7037 CGAGAGCGGCTCTTGGAAAAACAGCTAGTATTTAAGAGCTTGACCAACCGAGGTAGGA 7096
Qy 2344 ValValThrPheSerCysLysGlyGlyHisValLeuGlnGlyProSerValLeuLysCys 2363
Db 7097 GTTGTGACATTTCTGTAAAGAGGCGATGCTCTGCAAGGGCCCTCTGCTCTCTGAAATGC 7156
Qy 2364 LeuProSerGlnGlnTrpAsnAspSerPheProValCysLeuIleValLeuCysThrPro 2383
Db 7157 TTGCCATCCCAAGCAATGAATGACTCTTTCCCTGTGTGTGAAGATGTTCTTTGTATCCCCA 7216
Qy 2384 ProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGlySerThr 2403
Db 7217 CCTCCCTTAATTTCTTTGGTGTCCCATTCCTCTCTGCTCTTTCATTTTGGAGTACT 7276
Qy 2404 ValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGln 2423
Db 7277 GTCAAGTATTTCTGTGTAGTGGGTTTTCTTAAGAGGAAATTTCTACCAACCTCTGCCAA 7336
Qy 2424 ProAspGlyThrTrpSerProLeuProGluCysValProValGluCysProGlnPro 2443
Db 7337 CCTGTATGGACCTGGAGCTCTCCACTGCCAGATGTGTTCAGTAGATGTCTCCAGCAGCTCTC 7396
Qy 2444 GluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeu 2463
Db 7397 GAGGAAATCCCAATGGAATCATTCATGTGCAAGGCTTGCCTATCTCAGCAGCAGCTCTC 7456
Qy 2464 TyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeuCysGlyGluAsn 2483
Db 7457 TATACCTCGCAAGCCAGGCTTTGAATTTGGTGGGAAATATCTACCAACCTTTGTGTGGAGAAAT 7516
Qy 2484 GlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLysProLysGlu 2503
Db 7517 GGTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTGAGTGCCTGAAACCCAGGAG 7576
Qy 2504 IleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSer 2523
Db 7577 ATTTGATGGCAATTTCTTTACAGGACCTTACACTATGGACAGACCGTTACTACTCT 7636
Qy 2524 CysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAsp 2543
Db 7637 TGCAACCGAGGCTTTCCGCTCGAAGGTCCTCAGTGCCTGACCTGTTAGACAGAGTGAT 7696

Db 263 CTGGCGCAGCGCTTC-----CGCGTCGCGTCTGTCGGGAGCTCAGCGAGCGCTGGAG 316
Qy 84 LeuValPheLeuValAspSerSerValGlyGluValAanPheArgSerGluLeu 103
Db 317 CTGTCTCTCTGTTGGATGATTCGTCCAGCGTGGCGAGTCAACTTCGACGCGAGCTC 376
Qy 104 MetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db 377 ATGTTGTCGCGCAGCTGCTCCGACTCTCCCGTGGTGGCCACGCGCACGCGGTGGCC 436
Qy 124 IleValThrPheSerSerLysSerValValProArgValAspTyrLleSerThrArg 143
Db 437 ATGCTGACCTTCCTGCTCAAGAACTACGTGTGGTGGCGCGTGGATTAATCTCCACCCGC 496
Qy 144 ArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluLeuProAlaLleSerTyrArg 163
Db 497 CGCGCGCGCCAGCAGTGGCGCTCTCTCCAGAGATCCCTGCCATCTCTCTACCGA 556
Qy 164 GlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnlleLeuLeuHisAla 183
Db 557 GGTGGCGGCACCTACACCAAGGCGCTCTCCAGCAGCGCGCAAAATCTTCTTCATGCT 616
Qy 184 ArgGluAsnSerThrLysValValPheLeuLleThrAspGlyTyrSerAsnGlyGlyAsp 203
Db 617 AGAGAAACTCAACAAAGTTGTAATTTCTCATCACTGATGATTAATTCATGGGGGAGAC 676
Qy 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluLlePheThrPheGlyIle 223
Db 677 CCTAGACCAATTCAGCGGTCTCTCGAGATTCAGAGTGGAGATCTTCACCTTTGGCATA 736
Qy 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
Db 737 TGGCAAGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAGGAGCACTGT 796
Qy 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaLeuAlaLeuHisGluAsp 263
Db 797 TACCTGCTACACAGTTTGAAGAAATTTGAGCGTTTAGCTGCGCGGCAATTCGATGAAGAT 856
Qy 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeuCysAspGlu 283
Db 857 CTACCTTCTGGAGTTTATTCAGATGATATGTCCTCACTCTCATATCTTTGTGATGAG 916
Qy 284 GlyIysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303
Db 917 GGCAGAGACTGCTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACACAGGCGCAATTT 976
Qy 304 GluCysIleCysGluLysGlyTyrTyrGlyGlyGlyLeuGlnTyrGluCysThrAlaCys 323
Db 977 GAGTGCATCTGTGAAGAGGGGTATTACGGGAAAGGTCTGAGTATGTAATGCACAGCTTGC 1036
Qy 324 ProSerGlyThrTyrLysProGluGlySerProGlyLysSerGlyIleSerSerCysIleProCys 343
Db 1037 CCATGGGGGACATACAACTGAGCTTCACGAGGAGATCAGCACTTCGATTCATGT 1096
Qy 344 ProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCysValCysArg 363
Db 1097 CCGGATGAAATACACACTCTCCACCTGGAAGCACATCCCTGGAAGACTGTGTCTGCAGA 1156
Qy 364 GluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
Db 1157 GAGGGATACAGGGGCACTGGCCAGACTGTGAATCTGTCCACTGCCCTGCCCTGAGCCCT 1216
Qy 384 ProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGly 403
Db 1217 CCCGAAATGGTTACTTTATCCAAAACACTTGCACCAACCACTTCATGACGCTGTGGG 1276
Qy 404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCysLeuProAsn 423
Db 1277 GTCCGATGTCACTGGATTTGATCTTGTGGGAGGAGCATCATCTTATGCTTACCCCAAT 1336
Qy 424 GlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGln 443

Db 1337 GGTTTGTGGTCCGGTTCAGAGAGCTACTGCAGAGTAAGAACATGTCCTCATCTCCGCCAG 1396
Qy 444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeu 463
Db 1397 CCGAAACATATGCCACATCAGCTGTTCTACAGGGGAAATGTTATATAGAACACATGTTTG 1456
Qy 464 ValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGlyAan 483
Db 1457 GTTGCTGTGATGAAGGGTACAGACTAGAAGCGAGTATAGCTTACTTGTTCAGAGAAAC 1516
Qy 484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
Db 1517 AGCCAGTGGGATGGGCCAGAACCCCGGTGTGTGGAGCGCCACTGTTCACCTTTCCAGATG 1576
Qy 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
Db 1577 CCCAAGATGTCAATATCCCCCAACTGTGGCAAGCAGCCAGCCCAAAATTTGGGAGC 1636
Qy 524 IleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
Db 1637 ATCTGCTATGTAGTTGCCCAAGGGTTCAATTTATCTGGAGTCAAAAGAAATGCTGAGA 1696
Qy 544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAspValGlu 563
Db 1697 TGTACCACTTCTGGAAATGGAAATGTGGAGTTCAGGCAGCTGTGTGTAAAGACGTGGAG 1756
Qy 564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnAspSer 583
Db 1757 GCTCTCAAAATCAACTGTCTTAAGACATAGAGCTAAGACTCTGGAAACAGCAAGATTTCT 1816
Qy 584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerVal 603
Db 1817 GCCAATGTGTACCTGGCAGATTCCAACAGCTAAAGCAACTCTGGTGAAGAGGTGCAGTC 1876
Qy 604 HisValHisProAlaPheThrProTyrLeuPheProIleGlyAspValAlaIleVal 623
Db 1877 CGCGTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGTTGCTATCGTA 1936
Qy 624 TyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
Db 1937 TACACGGCAACTGACCTATCCGGCAACCCAGGCCAGCTGCATTTTCCATATCAAGGTTAT 1996
Qy 644 AspAlaGluProProValIleAspTrpCysArgSerProProValGlnValSerGlu 663
Db 1997 GATGCAAGACCACTGTGCATAGACTGGTGCAGATCTCCACCTCCGCTCAGGTCTGGAG 2056
Qy 664 LysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyValAlaGluLeu 683
Db 2057 AAGTACATGGCGCAAGCTGGATGAGCTCAGTTCTCAGACAACTCAGGGGCTGAAATTG 2116
Qy 684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
Db 2117 GTCAATACCAAGACTATACACAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAG 2176
Qy 704 TyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLys 723
Db 2177 TATACGCCACTGACCCCTCAGGTAAACAGGATATGTGATATCCATATTTGTGATGAA 2236
Qy 724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
Db 2237 GGTTCCTCCTGTGAATTTCCATTCCACACTGTGAATGGGATTTTATATGCACCTCCAGT 2296
Qy 744 AsnThrGlyValAanCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
Db 2297 AATACTGGAGTCACTGTACATTAACCTTGTGTGAGGGCTACGATTTTCCAGAGAGGGTCT 2356
Qy 764 ThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrThrGlu 783
Db 2357 ACTGACAAAGTATTAATGTGCTTATGAAGATGGCGTCTGGAAACCAACATATATACCACTGA 2416
Qy 784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
Db 2417 TGGCCAGACTGTGCAAAAACOGTTTTCGAAACCAACCGGGTTCAAGTCTCTTTGAGATGTTT 2476

Qy 804 TyrIysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu 823
Db 2477 TACAAAGCAGCTCGTGTGTGATGACTCAGATCTGATGGAAGAAGTTTCTGGAAGCAATTTGAG 2536
Qy 824 ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
Db 2537 ACGACCTTGGGGAANAATGGTCCCATCATTTGTAGTGTATGACGAGACATTTGACTGCAGA 2596
Qy 844 LeuGluGluAenLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly 863
Db 2597 CTGGAGGAGAACCTTGACCAAAAAATATTGCTCCTAGAAATATAATTAATGATATGCAATGCAAAATGGC 2656
Qy 864 PheAlaIleGlyProGlyGlyTyrGlyValAlaAenArgLeuAspTyrSerTyrAspAsp 883
Db 2657 TTTGCAATTTGGTCCAGGTGGCTGGGTGCAGCTAATAGCTGGATTAATCTCTTACGATGAC 2716
Qy 884 PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle 903
Db 2717 TTCCTGGACACTGTGCAGAAACACGCCACGATCGGCATGCCAGTCTCACGGATT 2776
Qy 904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
Db 2777 AAAAGAGTGGCCCATTTATCTGACTATAAAATTAAGTTAAATTTTAAACATCACAGCTAGT 2836
Qy 924 ValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnArgLeuLeu 943
Db 2837 GTGCCATTTACCCGATGGAAGAAATGATACCTTGAATGGGAAATACGACCAAGCTCCTT 2896
Qy 944 GlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyr 963
Db 2897 CAGACATTTGGAAACTATCACAAATAACTGAAAGAGACTCTCAACAGACCCCATGTAT 2956
Qy 964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLys 983
Db 2957 TCCCTTCAGCTTGCAATCAGAAATACTTATAGCCGACAGCAATTTCAATAGAACAAAAAG 3016
Qy 984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003
Db 3017 GCTTCCCTCCCTCTGAGACACAGGCTCAGTGTGAGAGGGCGTATGTGTCAATTTGCCCT 3076
Qy 1004 LeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr 1023
Db 3077 TTGGGAACCTATTAATCTGGAAACATTTCACTGTGAAAGCTGCGGATCGGATCCTAT 3136
Qy 1024 GlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyr 1043
Db 3137 CAAGATGAAGAGGGCAACTTGATGTCAGACTTTTGCCTCTCTGGGATGTACACGGAAAT 3196
Qy 1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
Db 3197 ATCCATTCAGAAACATCTCTGATTTGAAGCTCAGTGTAAACAGGACCTACTCTATAC 3256
Qy 1064 SerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySer 1083
Db 3257 AGTGGACTTGAGACTGTGAATCGTGTCCACTGGGCACTTATCAGCAAAAATTTGGTTCC 3316
Qy 1084 ArgSerCysLeuSerCysProGluAenThrSerThrValLysArgGlyAlaValAsnIle 1103
Db 3317 CGGAGCTGCCTCTCGTGTCCAGAAAACCTCAACTGTGAAAAGAGGAGCGCGTGAACATT 3376
Qy 1104 SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys 1123
Db 3377 TCTGCATGTGGAGTTCTCTTGTCCAGAAAGAAATTTCTCGGTTCTGGGTAAATGCCCTGT 3436
Qy 1124 HisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys 1143
Db 3437 CACCCATGTCCTGGTACATTTACCAACTAATAGCAGGAGAGGCTTCTCGCTGGCCTGT 3496
Qy 1144 ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPhe 1163
Db 3497 CCCTTTTATGGAACCTACCCCAATTCCTGCTTCCAGATCCATCACAGAAATGTTCAAGTTT 3556

Qy 1164 SerSerThrPheSerAlaAlaGluSerValValProProAlaSerLeuGlyHisIle 1183
Db 3557 AGTTCACTTTCTCAGCGCAGAGAAAGTGTGGTCCCTCTCTCTTGGACATATT 3616
Qy 1184 LysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHis 1203
Db 3617 AAAAAGAGGCATGAATCAGCAGTCAAGTCAATGAATGCTTCTTTAACTTCGCCAC 3676
Qy 1204 AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr 1223
Db 3677 AATAGTGGAACTCGCCGCAACTTGGGCGTGGTGTATGTTGTCTCTGTCCACTTGGATAT 3736
Qy 1224 ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAenAsn 1243
Db 3737 ACAGCTTTAAGTGTGAACACAGACATCGATGAGTGCAGCCCACTGCTTGCCTCAACAT 3796
Qy 1244 GlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGly 1263
Db 3797 GGAGTTTGTAAAGACCTAGTTGGGGAATTCATTTGTAGTGGCCATCAGGTTACACAGT 3856
Qy 1264 GlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAenLysGlyIle 1283
Db 3857 AAGCACTGTGAATTCGAACATCAATCAATGATGTCAGTCTAATCCATGTAGAAATCAGGCCACC 3916
Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
Db 3917 TGTGTGGATGAATTAATTCATACAGTGTGAATGTGACCCAGGATTTTCAGGCAAAAGG 3976
Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
Db 3977 TGTGAACACAGGTATGATCACTCAGTGTATTAATTAATTAACCTTAATTAATGCACTGTGAA 4036
Qy 1324 AspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCysGly 1343
Db 4037 GACCAGGTGGGGGAATCTTGTGCAATGCCCACCTGGGATTTTGGGTACCCCATGTGGA 4096
Qy 1344 LysAsnValAspGlyCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
Db 4097 AAGAACGTTCGATGAGTGTCTCAGTCAGCATTCGCAAAATGAGAGCTACTGTAAAGCGT 4156
Qy 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
Db 4157 GCCAATAGCTTCAGGTGCTGTGTGTCAGCTGAGCATTCGCAAAATGAGAGCTACTGTAAAGCGT 4216
Qy 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
Db 4217 ATCAATGAATGTCACTTAATCCATGTAGAAATCAGGCCACCTGTGTGGATGAATTAAT 4276
Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
Db 4277 TCATACAGTTGTAAATGTCCAGCAGGATTTTCAGGCAAAAGGTGTGAAACAGAACAGTCT 4336
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
Db 4337 ACAGGCTTTAAACCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCTAGTAGTGC 4396
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSerSerAspAspMet 1463
Db 4397 ATGCTCCCATCTCTCCATGCTCTAACCCTGACTTCTGGATGAATTCCTCTGACGACATG 4456
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuLeu 1483
Db 4457 AACTATGGAACACCAATCTCTATGTCAGTTGATACGGCAGCGCAATATCTTGTCTCTG 4516
Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
Db 4517 ACTGATTAATAACGGGTGGGTCTCTTTATGTGAATGGCAGGAAAGATTAACAAATCTGCC 4576
Qy 1504 SerValAsnAspGlyValArgTyrHisHisIleAlaIleThrTyrThrSerAlaAsnGlyIle 1523
Db 4577 TCGGTGAATGATGGCGATGGCATCATATTGCAATCATTGGAACAGTACTGTGTGGAGCC 4636
Qy 1524 TrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu 1543

Db 4637 TGGAGGGTCTATATAAATGGGAATTTATCTGACGGTGGTACTGGCTCTCCATTGGCAAA 4696
QY 1544 ProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLysGlyGly 1563
Db 4697 GCCATACCTGGTGGCGGTGCTAGTTAGTTCTTTGGGCAAGAGCAGACAAAGAGAGGGG 4756
QY 1564 PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal 1583
Db 4757 TTCACCCGGCTGAGTCTTTTGTGGGCTCCATTAAGCCAGCTCAACCTCTGGGACTATGTC 4816
QY 1584 LeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGly 1603
Db 4817 CTGTCTCCACAGCAGGTGAAGTCACTGGCTTACCTCTCCAGGGAATCAGTAAGGA 4876
QY 1604 AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLysIleAspSer 1623
Db 4877 AACGTGTAGCATGGCTGATTTCTTGTTCAGGAATTTGTGGGGAAGTGAAGATCGATTCT 4936
QY 1624 LysSerIlePheCysSerAspCysProAtqLeuGlyGlySerValProHisLeuAtqThr 1643
Db 4937 AAGAGCATATTTTGTTCAGATGCCACGCTTGGGAGGGTCACTGCTCATCTGAGAACT 4996
QY 1644 AlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln 1663
Db 4997 GCATCTGAAGATTTAAACCCAGGTTCCAAGTCAATCTGTTCTGTGAACAGGCTCCAG 5056
QY 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuPro 1683
Db 5057 CTGGTCGGGAACCTCTGTCAGTACTGTCTGAATCAGGACAGTGGACACCAACCACTCCCC 5116
QY 1684 HisCysGluAspGlySerCysGlyValProProLeuGluAsnGlyPheHisSerAla 1703
Db 5117 CACTGTGAAGCGCATCTGCTGGGAGTCCACCTCTTTGGAGAAATGGCTTCCATTCAGCC 5176
QY 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu 1723
Db 5177 GATGACTTCTATGCTGGCAGCACAGTAACCTACCAGTGCACCAATGGCTACTATATTG 5236
QY 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys 1743
Db 5237 GGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGCGCTTCCACCATCTGC 5296
QY 1744 LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn 1763
Db 5297 TTAGATGTCTGATGTCAGTGTGATGATCAGATTGTAGTGAAGCATGCTTCTTGCTGAAC 5356
QY 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCys 1783
Db 5357 GTAGATGGATCCTACATATGTTTATGTGTCACCGTACACAGGAGATGGGAAAACCTGT 5416
QY 1784 AlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu 1803
Db 5417 GCAGAACCTATAAAATGTAAGGCTCCAGGAAATCCGGAAAATGGCCACTCTCAGGTGAG 5476
QY 1804 IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly 1823
Db 5477 ATTTATACAGTAGTGGCCGAAGTCACATTTTGTGTGTCAGGAGGATACCACTGATGGGA 5536
QY 1824 ValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyrCysLys 1843
Db 5537 GTAAACCAATACATGTTTGGAGTCTGGAGATGGGAATCATCTAATACCATATTGTAAA 5596
QY 1844 AlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGluLeuAlaPhe 1863
Db 5597 GCTGTTTCATGTGTTAAACCGGCTATTCCAGAAAATGGTTGCAATTGAGGAGTTAGCAATT 5656
QY 1864 ThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLys 1883
Db 5657 ACTTTTGGCAGCAAGTGAATATAGTGTAAATGAAGGATATATCTTGGCCGGTGATPAA 5716
QY 1884 GluSerSerCysLeuAlaAsnSerSerTrpHisSerGluSerProValCysGluProVal 1903

Db 5717 GAATCATCTCTGTCTGTAAACAGTTCTTGAGAGTCAATTCCTCCTGTGTGTGAACCACTG 5776
QY 1904 LysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeuThrTyr 1923
Db 5777 AAGTGTCTTAGTCCGGAATAATATAATATGAATAATATATTTTGAATGGGCTTACCTAC 5836
QY 1924 LeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIle 1943
Db 5837 CTTTCTACTGCATCATATTCTATCATAGGATACAGCTTACAGGGGCCCTTCCATTATT 5896
QY 1944 GluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCys 1963
Db 5897 GAATGCAAGCTTCTGGCATCTGGGACAGAGCGCACCTGCTGTACCTGCTCTCTCTGT 5956
QY 1964 GlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsn 1983
Db 5957 GGAGAACCACTGTCATCAAGATGCTGTCTAAGGGGAATAACTTCACTTTCAGGAAC 6016
QY 1984 ThrValThrTyrThrCysLysGlyGlyTyrThrLeuAlaGlyLeuAspThrIleGluCys 2003
Db 6017 ACCGTCACTTACACTTGCACAAAGAGGCTATCTTCTGTGGTCTTGACACCATTAATGC 6076
QY 2004 LeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCysAspGlu 2023
Db 6077 CTGGCCGACGCAAGTGGAGTAGAAGTGAACCCAGCAGTGCCTGGCTGTCTCTGTGTATGAG 6136
QY 2024 ProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAspIleAla 2043
Db 6137 CCACCAATTGGACCAACGCTCTCCAGAGACTGCCATCGGCTCTTTTGGAGACATTGCA 6196
QY 2044 PheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAla 2063
Db 6197 TTTCTACTGCTCTGTAGTGTACAGCTAGCAGACAAATCCCAGCTTCTCTGCAATGCC 6256
QY 2064 GlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAlaHisPheCys 2083
Db 6257 CAGGGCAAGTGGGTACCCCCAGAGGTCAGACATGCCCGCTTGTATAGCTCATTTCTGT 6316
QY 2084 GluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAlaLysPheAla 2103
Db 6317 GAAACCTCCATCGGTTTCTATAGCATCTTGGAACTCTGGAACTGTGAGCAAGCAAAATTTGCA 6376
QY 2104 AlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLys 2123
Db 6377 GCTGGCTCAGTGTGACCTTTAAATGCATGGAAGGCTTTGTACTGAACACCTCAGCAAG 6436
QY 2124 IleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGlnCysIlePro 2143
Db 6437 ATTGAATGTATGAGAGTGGGCAAGTGGAAACCTTTCCTCCCATGTCCATCCAGTGCATCCT 6496
QY 2144 ValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSer 2163
Db 6497 GTGGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCAAACTACAGT 6556
QY 2164 PheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyLysLys 2183
Db 6557 TTTGGACCATGTGGTGTACAGCTGCACAGGGGTTCTACATCAAGAGGGGAAAGAG 6616
QY 2184 SerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisProValSer 2203
Db 6617 AGCACCTGCCAAGCCACAGGCGAGTGGAGTAGTCTTATACCGAGCTGCCACCGGTATCT 6676
QY 2204 CysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePhe 2223
Db 6677 TGTGGTGAACCACTTAAGGTGAGAAATGGCTTTCTGGAGCATACAACTGGCAGGATCTTT 6736
QY 2224 GluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPhe 2243
Db 6737 GAGGTGAAGTGAAGTATCAGTGTAAACCCGGCTATAAGTCACTGCGGAAGTCTCTGATTT 6796
QY 2244 ValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValProLeuAsp 2263
Db 6797 GTCTGCCAAGCCAATCGCCACTGGCAGATGAATCCCTCTGTGATGTGTGTCTCTCTCGAC 6856

2264 CysGlyLysProProPheGlnGlnGlyPheMetLysGlyGluAsnPheGluValGly 2283
6857 TGTGGAAACCTCCGCCATCCAGATGCTTTCATGAAGAGGAAACCTTTGAAGTAGGG 6916
2284 SerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerTrpThr 2303
6917 TCCAAAGTTTCAGTTTTCTGTAAAGGGTTATAGCTTTGTGGTACAGTTCTTTGGACA 6976
2304 CysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAlaLysCys 2323
6977 TGTCAAGAACTGGCAAAATGGAATAGAAAGTCAAAATCCAAAGTGCATGCTGCCAAGTGC 7036
2324 ProGluProProLeuLeuGlnGlnLeuValLeuLysGluLeuThrThrGluValGly 2343
7037 CCAGAGCCGCCCTCTTGGAAACCAACAGCTAGTAATTAAGAGAGTTGACCAACCGAGTAGGA 7096
2344 ValValThrPheSerCysLysGluGlyHisValLeuGlnGlnProSerValLeuLysCys 2363
7097 GTTGTGACATTTTCTGTAAAGAGGGCATGCTGCAAGGCCCTCTCTGTCTGTGAANTGC 7156
2364 LeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCysThrPro 2383
7157 TTGCCATCCAGCAATGGAATGACTCTTCCCTGTTGTGAAGATTGTTCTTTGTACCCCA 7216
2384 ProProLeuLysSerPheGlyValProLysProSerSerAlaLeuHisPheGlySerThr 2403
7217 CCTCCCTCAATTTCTTGTGCTGCCCATTCCTTCTCTCTCTTCAATTTTGAAGTACT 7276
2404 ValLysThrSerCysValGlyGlyPhePheLeuValGlyAsnSerThrThrLeuCysGln 2423
7277 GTCAAGTATCTGTGTAGTGGTGTGTTTCTTAAGAGGAATTTCTACCAACCTCTGCCAA 7336
2424 ProAspGlyThrTrpSerProLeuProGluCysValProValGluCysProGlnPro 2443
7337 CTGTATGGACCTGGAGCTCTCCACTGCCAGAAATGTTTCCAGTAGAATGTCCCAACT 7396
2444 GluGluLeuProAsnGlyIleLeuAspValGlnGlnGlyLeuAlaTyrLeuSerThrAlaLeu 2463
7397 GAGGAATCCCAATGGAATCATTTGATGTGCAAGGCTTTGCCCTATCTCAGCACAGCTCTC 7456
2464 TyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsn 2483
7457 TATACCTGCACAGCAGCTTTGATTTGGTGGGAAATACCTACCAACCTTTTGTGGAGAAAT 7516
2484 GlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLysProLysGlu 2503
7517 GGTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTTGAGTGCCTGAAACCCCAAGGAG 7576
2504 IleLeuAsnGlyLysPheSerThrAspLeuHisTyrGlyGlnThrValThrTyrSer 2523
7577 ATTTTGAATGGCAAAATTTCTTACACGGACCTACACTATGGACAGACCGTTACTCTCT 7636
2524 CysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAsp 2543
7637 TGCACCGAGGCTTTCCGCTCGAGGTCCAGTGCTTCCCTGATCTTGTAGACAGCTGAT 7696
2544 TrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProLysGlu 2563
7697 TGGATGTAGATGCCCATTTGCAATGCCATCTCACTGTGATTTCCCAACCAACCAATGAA 7756
2564 AsnGlyPheValGluGlyAlaAspTyrSerTyrGlyValIleIleTyrSerCysPhe 2583
7757 AATGGTTTTGTAGAGGTGCAGATTAACAGCTATGGTGCCATATCATCTACAGTTGCTCTC 7816
2584 ProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSer 2603
7817 CTGGGTTTTCAGTGGCTGGTCAATGCCATGCAGACCTGTGTAGAGAGTCAGGTCACT 7876
2604 SerIleProThrCysMetProLysAspCysGlyLeuProProHisIleAspPheGlyAsp 2623
7877 TCCATCCCAACATGTATGCCAATAGACTGTGGCTCTCCCTCTCATATAGATTTTGGAGAC 7936

2624 CysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspMetMetGluVal 2643
7937 TGTACTAAACTCAAGATGACCAGGATATTTTGACAGAGACGACATGATGGAAGTT 7996
2644 ProTyrValThrProHisProTyrHisLeuGlyValAlaLysThrTrpGluAsn 2663
7997 CCAATATGACTCTCTCACCCTCTTATCAATTTGGGAGCAGTGGCTAAACCTGGGAAAT 8056
2664 ThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyr 2683
8057 ACAAAAGAGTCTCTCTGCTACATTCATAACTTTCTGTATGTATACCATGGTTTCATAC 8116
2684 ThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValIleLeuCysGlnGluAspGly 2703
8117 ACTGTAAATCCAGGATATGAATCTCTGGGAAACCTCTGTCTGATCTGCCAGAGAGATGGA 8176
2704 ThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaPro 2723
8177 ACTTGAATGGCAGTGCACCATCTGCAATTTCAATTTGAATGTGACTTGCCTACTGCTCTC 8236
2724 GluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCys 2743
8237 GAAAATGGCTTTTTCGGTTTTTACAGAGACTAGCATGGGAAGTCTGTGCCAGTATAGCTGT 8296
2744 LysProGlyHisIleLeuAlaGlySerAspLeuValGlyLeuGluAsnArgLysTrp 2763
8297 AAACCTGGACACATTTCTAGCAGCTCTGACTTAAAGCTTTGTCTAGAGATAGAAAGTGG 8356
2764 SerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnProValMetAsn 2783
8357 AGTGTGTGCTCTCCACGCTGTGAAGCCAATTTCAATGCAAAAGCAAAATCCAGTCATGAAT 8416
2784 GlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspPro 2803
8417 GGATCCATCAAAAGGAAGCAACTACATACCTTGAGCAGCGTTGTATGATGTGACCCCTC 8476
2804 GlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrpAspGlu 2823
8477 GGATATGTCTGTATGGCACTGAGAGAGACATGTCAGGATGACAAAACTGGGATGAG 8536
2824 AspGluProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyGln 2843
8537 GATGAGCCCATTTGCAATCTCTGTGACTGCAGTTTCACTCCCTCAGCTCAGCAATGGCCAG 8596
2844 ValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPhe 2863
8597 GTGAGAGGAGACGAGTACACATTTCCAAAGAGATTTGAATACACTTGCATTAAGAGGGTTC 8556
2864 LeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThr 2883
8657 TTGCTTGAAGGAGCCAGAGTGGGTTTCTTGTCCCAATGGAAGTTGGAGTGGAGCCACT 8716
2884 ProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGlu 2903
8717 CCGACTGTGTGCTGTCTAGATGTGTCACCCCGCCCACTGGCCCAATGGGGTGACCGAA 8776
2904 GlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeu 2923
8777 GGCCTGGACTATGGCTTTCATGAGGAAGTAACATTTCACTTGTCCAGGGGCTACATCTTG 8836
2924 HisGlyValaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeu 2943
8837 CACGGTGTCTCAAAACTCACTCTGCTCAGTGCAGATGGCACTGGGATGCAGAGATTCCTCTC 8896
2944 CysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPhe 2963
8897 TGTAAACCACTCAACTGTGGACCTCTGAAGATCTTGGCCCAATGGTTTCCCTAATGGTTTT 8956
2964 SerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGly 2983
8957 TCTTTTATTTATGGGGGCCATATACAGTATCAGTGTCTTCTCTGTTATTAAGCTCCATGGA 9016
2984 AsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerProSerCysLeu 3003

Db 9017 AATTCATCAAGAGGTGCTCTCCAAATGGCTCTGGAGTGGCAGCTCACCTTCTGCTG 9076
Qy 3004 ProCysArgCysSerThrProValIleGluThrValAsnGlyThrAspPheAsp 3023
Db 9077 CCTTGCAGATGTTCCACACCAAGTAATTAATGAAATATGGAATCTCAATGGGACAGATTTTGAC 9136
Qy 3024 CysGlyValAlaAATGileGlyCysPheLysGlyPheLysLeuLeuGlyLeuSerGlu 3043
Db 9137 TGTGGAAGGAGGCGCGGATTCAGTGCCTTCAAGGCTTCAAGCTCTTAGGACTTTCGAA 9196
Qy 3044 IleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSer 3063
Db 9197 ATCACTGTGAAGCGATGGCAGTGGAGCTCTGGTTCCTCCCACTGTGAACACACTTCT 9256
Qy 3064 CysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGlu 3083
Db 9257 TGTGCTTCTCTTCCAAATGATACCAATGCGTTCATCAGTGAGACAGCTCTTGGAAAGGA 9316
Qy 3084 AsnValIleThrThrSerCysArgSerGlyThrValIleGlnGlySerSerAspLeuIle 3103
Db 9317 AATGTGATAACTTACAGCTGCGAGTCTGGATATGTCATACCAAGGCGATTCAGATCTGATT 9376
Qy 3104 CysThrGluLysGlyValTrpSerGlnProValCysGluProLeuSerCysGly 3123
Db 9377 TGTACAGAAAGGGGTATGAGCCAGCTTATCCAGTCTGTGAGCCCTGTCTGTGG 9436
Qy 3124 SerProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrThrGluSerGlu 3143
Db 9437 TCCCCACCGTCTGTGCGCAATGCAAGTGCACCTGGAGAGGCACACCTATATGAAGTGA 9496
Qy 3144 ValLysLeuArgCysLeuGluGlyThrMetAspThrAspThrAspThrPheThrCys 3163
Db 9497 GTGAAACTCAGATGTCTGGAAGGTATACGATGATACGATGATACAGATACATATCAGCTGT 9556
Qy 3164 GlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeu 3183
Db 9557 CAGAAAGATGTCGTGCTGCTTCCCTGAGAGATCTCTGAGATCTCTTAAAAATGCTCTC 9616
Qy 3184 ProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnVal 3203
Db 9617 CCGGAAACATAACACATATATCTGTATGCGGACGATTTCAAGTGTGAATAGCAAGTT 9676
Qy 3204 SerValSerCysAlaGluGlyThrPheGluGlyValAlaAsnIleSerValCysGlnLeu 3223
Db 9677 TCTGTGTCATGTGCAGAGGGTATACCTTTGAGGGAGTTAAATATCATATCATATGTCAGCTT 9736
Qy 3224 AspGlyThrTrpGluProPheSerAspGluSerCysSerProValSerCysGlyLys 3243
Db 9737 GATGGAACCTGGAGGCCACCATCTCCGATGATCTTGCAGTCCAGTCTTCTTGTGGGAA 9796
Qy 3244 ProGluSerProGluHisGlyPheValValGlySerLysThrThrPheGluSerThrIle 3263
Db 9797 CCTGAAAGTCCAGACATGGATTTGTGGTGGCAGTAAATAACCTTTGAAAGACACAAAT 9856
Qy 3264 IleThrGlnCysGluProGlyThrGluLeuGluGlyAsnArgGluArgValCysGlnGlu 3283
Db 9857 ATTTATCATGTGAGCCCTGGCTATGAACTAGAGGGGACAGGAAACGCTGTCTGCCAGAG 9916
Qy 3284 AsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeu 3303
Db 9917 AACAGACAGTGGAGTGGGGGTGGCAATATGCAAAAGACAGAGTGTGAATCCACTT 9976
Qy 3304 GluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnValValThr 3323
Db 9977 GAATTTCTCAATGGGAAGCTGACATTTGAAACACAGACGACTGGACCCCAACGCTGGTATAT 10036
Qy 3324 SerCysAsnArgGlyThrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGly 3343
Db 10037 TCCTTCACACAGAGGCTACAGTCTTGAAGGGGCCATCTGAGGGGCACCTGCACAGAAATGA 10096
Qy 3344 ThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPheValIle 3363

Db 10097 ACCTGGAGCCACCAGCTCCCTCTCTGCAAAACCAATCCATGCCCTGCTCTTTTGAT 10156
Qy 3364 ProGluAsnAlaLeuLeuSerGluLysGluPheThrValAspGlnAsnValSerIleLys 3383
Db 10157 CCGGAAATGCTCTGCTCTGAAAGAGTTTATGTTGATCAGATGTGTCATCAAA 10216
Qy 3384 CysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThr 3403
Db 10217 TGTAGGAAGGTTTCTGCTGAGGGCCAGCGCATCATTTACCTGCAACCCCGACGAGCG 10276
Qy 3404 TrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisValGlu 3423
Db 10277 TGGACACAGACAGCGCAAAATGTGAAAAATCTATGTGGTCCACAGCTCAGCTAGAA 10336
Qy 3424 AsnAlaIleAlaArgGlyValHisThrGlnTrpGlyAspMetIleThrTrpSerCysThr 3443
Db 10337 AATGCAATGTCTGAGGGGTACATTTATCATATGAGACATCATCACCTACTCATGTTAC 10396
Qy 3444 SerGlyThrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThr 3463
Db 10397 AGTGATACATGTTGGAGGTTTCTTGAGAGTGTGTTTGTAGAAAAATGGAACATGACA 10456
Qy 3464 SerProProfileCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGln 3483
Db 10457 TCACCTCTATTTGGCAGAGCTGTCTGTGCAATTTCCATGTGAGATGGGGGCATCTGCCAA 10516
Qy 3484 ArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGluProIle 3503
Db 10517 CGCCCAATGCTGTCTGTCTGAGAGGGCTGAGTGGGGCGCTCTGTGAGAACCAATC 10576
Qy 3504 CysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProThrGlnCysAspCysPro 3523
Db 10577 TGCATCTCTCTGCTGCTGAGAGGCTGCTGTGTGGCCCTTACCAGTGTGACTGCCCG 10636
Qy 3524 ProGlyThrThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGly 3543
Db 10637 CCTGCTGGAGCGGCTCTCGCTGTATACAGCTGTGTGCGAGTCTCCCTGCTTAAATGGT 10696
Qy 3544 GlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCys 3563
Db 10697 GGAAATGTGTAAGACCAACGATGTCACTGTCTTCTTCTTGGACGGGACATAACTGT 10756
Qy 3564 SerArgLysArgArgThrGlyPhe 3571
Db 10757 TCCAGAAAAGGAGGACTGGGTTT 10780

RESULT 7

US-10-028-248A-5

; Sequence 5, Application US/10028248A

; Publication No. US20030235882A1

; GENERAL INFORMATION:

; APPLICANT: Shimkete, Richard

; APPLICANT: Patturajan, Meera

; APPLICANT: Vernet, Corine

; APPLICANT: Casman, Stacie

; APPLICANT: Malyankar, Uriel

; APPLICANT: Shenoy, Suresh

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gangolli, Esha

; APPLICANT: Miller, Charles

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Smithson, Glennda

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Liu, Xiaohong

; APPLICANT: Colman, Steven

; APPLICANT: Tchernev, Velizar

; APPLICANT: Si, Jingsheng

; APPLICANT: Edinger, Shlomit

; APPLICANT: Stone, David

; APPLICANT: Sciore, Paul

```

; APPLICANT: Millet, Isabelle
; APPLICANT: Rothenberg, Mark
; TITLE OF INVENTION: No. US20030235882a1el Nucleic Acids and Polypeptides and Methods
; TITLE OF INVENTION: Thereof
; PILE REFERENCE: 21402-222
;

```

Alignment Scores:		
Pred. No.:	0	11158
Score:	19566.00	13503
Percent Similarity:	98.57%	Conservative: 16
Best Local Similarity:	98.12%	Mismatches: 47
Query Match:	97.96%	Indels: 4
DB:	16	Gaps: 2

1697 CTGAGATGACCACTTCTGGAAATGGAATGTCGGAGTTCAGCAGCTGTGTGAAGAC 1756
562 ValGluAlaProGlnIleAenCysProLysAspIleGluAlaLysThrLeuGluGlnGln 581
1757 GTGAGGCTCTCTCAATCAACTGCTCTAAGCAGCATAGAGGCTAAGACTCTGTGAACAGCAA 1816
582 AspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspIleSerGlyGluLysVal 601
1817 GATTCGTGCCAATGTACCTTGGCAGATTCACACAGCTTAAAGACAACTCTGGTGAAGAGGTG 1876
602 SerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAla 621
1877 TCAGTTCACAGTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGTGCT 1936
622 IleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisLys 641
1937 ATCGTATACAGGCAACTGACCTATCCGCAACAGGCAAGCTGCATTTTCCATATCAAG 1996
642 ValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGlnVal 661
1997 GTTATTGATGCAGAACCACTGTCATAGACTGGTGCGAGATCTCCACCTCCCGTCCAGGTC 2056
662 SerGluLysValHisAlaLysTrpAspGluProGlnPheSerAspAsnSerGlyAla 681
2057 TCGGAGAGGTACTGCGCCGACCTGGATGAGCCTCAGTTCTCAGAACAACTCAGGTGCT 2116
682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIle 701
2117 GAATTCGTCTATTACCAAGAGTATACACAGGAGAGCTTTTCCCTCAAGGGGAGACTATA 2176
702 ValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisLysVal 721
2177 GTACAGATATACAGCCACTGACCTCCCTCAGGCAATTAACAGGACATGTGATATTCATATGTC 2236
722 IleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
2237 ATAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATATGCACT 2296
742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
2297 CCAGATAATACCTGGAGTCAACTGATACATTAATTCCTGGAGGCTATGATTTTCACAGAA 2356
762 GlySerThrAspLysTyrCysAlaLysArgPheAlaAsnHisGlyPheLysSerPheGlu 781
2357 GGGTCTACTGACAAATATTATGTGCTTATGAGATGGCGTCTGGAAACCAACATATACC 2416
782 ThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
2417 ACTGAATGCCAGACTGTGCCAGTAAAGCTTTTGCAAACCAACCGGTTCAAGTCTTTGAG 2476
802 MetPheTyrLysAlaAlaArgCysAspThrAspLeuMetLysLysPheSerGluAla 821
2477 ATGTTCTACAAAGCAGCTCGTTGTGATGACACACAGATCTGATGAAGAAGTTTCTGAAGCA 2536
822 PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
2537 TTTGAGACGACCTGGGAAAATGGTCCCATCATTTTGTAGTGTGATGAGGAGGACATGAC 2596
842 CysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
2597 TGCAGACTGGAGGAGAACCTGACCAAAATATTTGCTAGATATTAATATGACTATGAA 2656
862 AsnGlyPheAlaIleGlyProGlyTrpGlyValAlaAlaAsnArgLeuAspTyrSerTyr 881
2657 AATGGCTTTGCAATTTGCTCAGGTGGCTGGGTGCAGCTAATAGCTGGATTAATCTTAC 2716
882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
2717 GATGACTTCTCGGACACTGTGCAAGAAACAGCCACCAAGCATCGGCAATGCGCAAGTCTCTCA 2776
902 ArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThr 921

2777 CGGATTAAGAAAGAGTGCCTCATTTATCTGACTATATAAAATTAAGTTAAATTTTAAACATCAC 2836
922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGlyArg 941
2837 GCTAGTGTGCCATTTACCCGATGAAGAAATGATACCTTGAATGGGAAAATTCAGCAACGA 2896
942 LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961
2897 CTCTCTCAGACATTTGGAACTATCAAAATAAATGAAGAGACTCTCAACAAGAGACCC 2956
962 MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
2957 ATGTATTTCTCTTTCAGCTTGCATCAGAAATACTTATAGCCGACAGCAATTTCAATTAGAAACA 3016
982 LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
3017 AAAAAGGCTTTCCCTTCTGCAGACCAAGCTCAGTGTGAGAGGCGGTATGTGTGTCAAT 3076
1002 CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
3077 TGCCCTTTGGAACTTATTAATCTGGAACATTTTCACTGTGAAGCTGCCGATCGGA 3136
1022 SerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyrThr 1041
3137 TCCTATCAAGATCAAGAGGCAACTTTGAGTGCAGCTTTGCCCTCTGGGATGTACAG 3196
1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
3197 GAATATATCCATTCAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAGGACCTTAC 3256
1062 SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPhe 1081
3257 TCATACAGTGGACTTGAGACTTGTGATCTGTGTCCACTGGGCACTTATCAGCCAAATTT 3316
1082 GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaVal 1101
3317 GGTTCCTCCGAGCTGCCTCTCTGTGTCCAGAAAACACCTCAACTGTGAAAAGAGGAGCCGTG 3376
1102 AsnIleSerHisCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMet 1121
3377 AACATTTCTGCATGTGGAGTCTCTGTCCAGAGAGGAAAATTTCTCGCGTTCTGGGTAAATG 3436
1122 ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeu 1141
3437 CCTGTCAACCATGTCTCTGTGACTATTACCACTTAATGACGAGGAGGCTTCTGTGCTG 3496
1142 AlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
3497 GCCTGTCCCTTTTATGGAACCTACCCCATTCGCTGTTCAGATCCATCAGAGATGTTC 3556
1162 SerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSerLeuGly 1181
3557 AGTTTTAGTTTCAACTTCTCAGCGCAGAGAAAGTGTGGTGGCCCTTCTCTCTTGA 3616
1182 HisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro 1201
3617 CATATTAAAGAGGAGCATGAAATCAGCATCAGCAGTCAATGATGCTTCTTTAACCTT 3676
1202 CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
3677 TGCCACAAATAGTGCACCTGCGCAGCAACTTGGGCGTGTGTTGTTGCTCTCTGTCCACTT 3736
1222 GlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
3737 GGATATACAGGTTTAAAGTGTGAACAGACATCGATGAGTGCAGCCCACTGCTTGGCTC 3796
1242 AsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
3797 AACAAATGGAGTTTGTAAAGACCTTAGTTGGGGAATTCATTTGTGAGTGCCTCATAGTTAC 3856
1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLys 1281
3857 ACAGGTAAGCAGCTGTGAATTTGAACATCAATGAATGTCACTTAATCCATGTAGAATACAG 3916

QY 1282 GlyIleCysValaspGlyValalaglyTyrArgCysThrCysVallysGlyPheValGly 1301
DB 3917 GCCACCTGTGTGGATGAATTAATTCATACAGTGTGTAATATGCAGCCAGGATTTTCAGGC 3976
QY 1302 LeuHisCysGluThrGluValAenGluCysGlnSerAsnProCysLeuAenAenAlaVal 1321
DB 3977 AAAAGGTGTGAACAGGTATGTATCACTCAGTGTATTATAAATCACTTAATTAATGCAGTC 4036
QY 1322 CysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArg 1341
DB 4037 TGTGAAGACCAAGTGTGGGGATTTCTGTGCAAAATGCCACCTGGATTTTGTGGTACCCCA 4096
QY 1342 CysGlyLysAenValaspGluCysLeuSerGlnProCysLeuAenGlyAlaThrCysLys 1361
DB 4097 TGTGAAGAACAGCTCGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGTCACTCTGTGAAA 4156
QY 1362 AspGlyAlaAenSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
DB 4157 GACGGTGCATAGCTTCAGGTGCCGTGTGTGAGCTGGCTTCACAGGATCACACTGTGAA 4216
QY 1382 LeuAenIleAenGluCysGlnSerAenProCysArgAsnGlnAlaThrCysValaspGlu 1401
DB 4217 TTGAACATCAATGAATGTCACTTAATCTAGTATGAGAAATCAGGCCACCTGTGTGGATGAA 4276
QY 1402 LeuAenSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
DB 4277 TTAATTCATACAGTGTGTAATGTGAGCAGGATTTTCAAGCAAAAGGTGTGAACAGAA 4336
QY 1422 GlnSerThrGlyPheAenLeuaspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
DB 4337 CAGTCTACAGGCTTTAACTCGATTTTGAGTTTCTGGCATCTAGGATATGTCATGCTA 4396
QY 1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAsp 1461
DB 4397 GATGGCATGCTCCCATCTCCATGCTTAACTGTACCTGTACCTTCTGGATGAAATCTCTGAC 4456
QY 1462 AspMetAenTyrGlyThrProIleSerTyrAlaValaspAenGlySerAenThrLeu 1481
DB 4457 GACATGAATATGGAACACCAATCTCCTATGAGTTGATACGGCAGGACCAATACCTTG 4516
QY 1482 LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValaspGlyArgGluLysIleThrAen 1501
DB 4517 CTCCTGACTGATTAACCGGTGGGTCTTTATGTGAATGGCAGGAAAGATACAAAC 4576
QY 1502 CysProSerValaspAspGlyArgTyrHisIleAlaIleThrTyrThrSerAlaAen 1521
DB 4577 TGTCCCTCGGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGCAAGTACTGGT 4636
QY 1522 GlyIleThrLysValTyrIleaspGlyLysLeuSeraspGlyValaglyLeuSerVal 1541
DB 4637 GGAGCTGGAGGGTCTATATAAATGGGAATTTATCTGAGCGGTGGTACTGGCCTCTCCATT 4696
QY 1542 GlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysGly 1561
DB 4697 GCAAGGCACTACTGGTGGCGGTGCAATTAGTTCTTGGCAAGAGCAAGCAAAAAGCA 4756
QY 1562 GluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAenLeuTrpAsp 1581
DB 4757 GAGGGGTTCACCCGGCTGAGTCTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGAC 4816
QY 1582 TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer 1601
DB 4817 TATGTCTGTCTCCACAGAGGTGAAGTCACTGGGTACTCTCTGCCCAGAGGAATCTAGT 4876
QY 1602 LysGlyAenValLeuAlaTrpProaspPheLeuSerGlyIleValGlyLysValLysIle 1621
DB 4877 AAAGGAACGTGTAGCATGGCTGATTTCTTGTGAGGAATTTGGGGAAGTGAAGATC 4936
QY 1622 AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeu 1641
DB 4937 GATTCTAAGAGCATATTTGTGTTCTGATTGCCACGCTTGGAGGGTCACTGCTCTCATCTG 4996

QY 1642 ArgThrAlaSerGluaspLeuLysProGlySerLysValAenLeuPheCysAspProGly 1661
DB 4997 AGACTGCTATCTGAAGATTTAAACCAAGTTTCCAAAGTCAATCTGTTCTGTGAACCGGC 5056
QY 1662 PheGlnLeuValGlyAsnProValGlnTyrCysLeuAenGlnGlyGlnTrpThrGlnPro 1681
DB 5057 TTTCCAGCTGGTCCGGAACCTGTGTCAGTACTGTCTGAATCAAGGACAGTGGACACAACCA 5116
QY 1682 LeuProHisCysGluaspGlySerCysGlyValProProProLeuGluAenGlyPheHis 1701
DB 5117 CTCGCCCATCTGTGAACGCAATTCGCTGTGGGGTGCACCTCTTTGAGGAATGCTTCCAT 5176
QY 1702 SerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAenAenGlyTyrTyr 1721
DB 5177 TCAGCGATGACTTCTATGCTGGCAGCAGTAACCTACAGTGCACAACTATGGCTACTAT 5236
QY 1722 LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAenGlyValSerPro 1741
DB 5237 CTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCGTTTCACCA 5296
QY 1742 SerCysLeuaspValaspGluCysAlaValGlySerAspCysSerGluHisAlaSerCys 1761
DB 5297 TCTCTGCTTAGATGTGCAATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGAT 5356
QY 1762 LeuAenValaspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLys 1781
DB 5357 CTGAACGTAGATGAGTCTCATATATGTTCTATGTTGCCACCGTACACAGGAGATGGAAA 5416
QY 1782 AenCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAenGlyHisSerSer 1801
DB 5417 AACTGTGCAGAACCTATATAATGTAAGGCTCCAGGAAATCCGGAATAATGGCCACTCTCA 5476
QY 1802 GlyGluIleTyrThrValGlyAlaValThrPheSerCysGlnGluGlyTyrGlnLeu 1821
DB 5477 GGTGAGATTTATACAGTGTAGTGTGCGAAGTCACTTTTCTGTGTGAGGAAGATACCACTG 5536
QY 1822 MetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisLeuIleProTyr 1841
DB 5537 ATGGAGTATACCAAAATCACTGTTTGGAGTCTGGAGATGGATCATCTAATACCATAT 5596
QY 1842 CysLysAlaValSerCysGlyLysProAlaIleProGluAenGlyCysIleGluGluLeu 1861
DB 5597 TGTAAAGCTGTTTCATGTGTGTAACCGCTATTCCAGAAATAATGGTTGCAATGGAGATTA 5656
QY 1862 AlaPheThrPheGlySerLysValThrTyrArgCysAenLysGlyTyrThrLeuAlaGly 1881
DB 5657 GCATTTACTTTTGGCAGCAAGTGCATATAGTGTGTAATAAGGATATATCTCTGGCCGT 5716
QY 1882 AspLysGluSerSerCysLeuAlaAenSerSerTrpSerHisSerProProValCysGlu 1901
DB 5717 GATAAAGATCATCTCTCTTCTTAACAGTCTCTGGAGTCAATCCCTCTCTGTGTGTGA 5776
QY 1902 ProValLysCysSerSerProGluAenIleAenAenGlyLysTyrIleLeuSerGlyLeu 1921
DB 5777 CCAGTGAAGTGTCTAGTCCGGAATAATAATAATGGAATAATATATATTTTGTAGTGGGCTT 5836
QY 1922 ThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer 1941
DB 5837 ACCTACCTTCTTACTGTCATCATATTCGCGATACAGGATACAGCTTACAGGCCCCCTTC 5896
QY 1942 IleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuVal 1961
DB 5897 ATTATTGAATGACCGCTTCTGTCATCTGGGACAGAGCGCCACTGCTGTCTCACTCGTC 5956
QY 1962 PheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAenAenPheThrPhe 1981
DB 5957 TTTCTGTGGAGAACCACTGCTCCATCAAGATGCTGTCTATACGGGGAATAATCTTCACTTTC 6016
QY 1982 ArgAenThrValThrTyrCysLysGlyTyrThrLeuAlaGlyLeuAspThrIle 2001
DB 6017 AGGAACCGCTCACTTACACTTGCAGGAAGGCTATCTCTTCTGTGTGTGTGTGTGTGTGT 6076
QY 2002 GluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCys 2021

6077 GAATGCTGCGCGAGCGAGTGGAGTAGAGTACCAGAGTGCCTGGCTGCTCTCTGT 6136
2022 AspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAsp 2041
6137 GATGAGCCACCAATGTCGACCAACGCGCTCTCCAGAGACTGCCCATCGGCTCTTTGGAGAC 6196
2042 IleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCys 2061
6197 AITGCAATCTACTACTGCTCTCTGATGGTTACAGCTTAGCAGACAAATTCACAGCTTCCTGC 6256
2062 AsnAlaGlnGlyLysTrpValProProGluGlnAspMetProArgCysIleAlaHis 2081
6257 AATGCCACGGCAGTGGGTACCCCGAGAGGTCGAAGACATGCCCGCTTGTATAGCTCAT 6316
2082 PheCysGluLysProProSerValSerTyrSerIleLeuGlnSerValSerLysAlaLys 2101
6317 TTTCTGTGAAAAACCTCATCGGTTTCTATAGCATCTTGGAAATCTGTGAGCAAGCAAAA 6376
2102 PheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSer 2121
6377 TTTGAGCTGGCTCAGTTGTAGCTTTAAATGATGGAAGGCTTTGTACTGAAACACTCA 6436
2122 AlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGlnCys 2141
6437 GCAGAGATTGAATGATGAGAGGTGGCGAGTGGAAACCTTCCCGCATGTCCATCCAGTGC 6496
2142 IleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsn 2161
6497 ATCCCTGTGCGGTGTGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCAAAAC 6556
2162 TyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyGlu 2181
6557 TACAGTTTTGGAGCCATGGTGGCTTACAGCTGCGCAACAAGGGGTTCATCAAGGGGAA 6616
2182 LysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisPro 2201
6617 AAGAGAGCACCTGCGAGGCCACAGCGCAGTGGAGTAGTCTCTATACCGAGTGCACCCG 6676
2202 ValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrGlyArg 2221
6677 GTATCTTGTGGTGAACCACTAAGGTTGAGATGGCTTTCTGAGAGATCAACTGGCAGG 6736
2222 IlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerPro 2241
6737 ATCTTTGAGAGTGAAGTGGATTCAGTGTAAACCGCGCTATAAGTCAGTGGGAAGTCTT 6796
2242 ValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValPro 2261
6797 GTATTTGTCTGCGACGCAATCGCCACTGGCAGTGGATGCCCTCTGTATGTGTCTCT 6856
2262 LeuAspCysGlyLysProProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGlu 2281
6857 CTCGACTGTGGAACCACTCCCGCATCCAGATGGCTTCATGAAGAGAGAAAACTTTGAA 6916
2282 ValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSer 2301
6917 GTAGGGTCCAAGGTTCAAGTTTCTGTAATGAGGGTTATGAGCTTGTGTGACAGTTCT 6976
2302 TrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAla 2321
6977 TGGACATGTCAGAAATCTGCGAAATGGAATAGAACTCAAAATCCAAAGTGCACTGCC 7036
2322 LysCysProGluProProLeuLeuGlnLeuValIleLysGluLeuThrThrGlu 2341
7037 AAGTGCCCAAGAGCGCGCTCTTGGAAAAACCAAGCTAGTATTAAGAGGTTGACCCAGAG 7096
2342 ValGlyValValThrPheSerCysLysGlyGlyHisValIleGlnGlyProSerValLeu 2361
7097 GTAGGAGTTGTGACATTTTCTCTGTAAGAGGGCATGTCTCGCAAGGCCCTCTGTCCG 7156
2362 LysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCys 2381

7157 AAATGCTTGCCATCCAGCAATGGAATGACTTTTCCCTGTTTGTGAAGATTGTTCTTGT 7216
2382 ThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGly 2401
7217 ACCCCACCTCCCTAAATTCCTTTGGTGTCCCATTCCTTCTCTGCTCTTCATTTTGA 7276
2402 SerThrValLysTyrSerCysValGlyGlyPheLeuArgGlyAsnSerThrThrLeu 2421
7277 AGTACTGTCAAGTATCTTGTGTAGGTGGGTTCCTTAGAGGAATCTTACCACCCCTC 7336
2422 CysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCysPro 2441
7337 TGCACACCTGATGGCACCTGAGCTTCCACTGCCAGATGTGTTCAGAGTAGAATGTCCC 7396
2442 GlnProGluLysProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThr 2461
7397 CAACCTGAGGAATCCCAATGGAATCAATGATGTGACAGCCCTTGCTATCTCAGCAACA 7456
2462 AlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGly 2481
7457 GCTCTCTATACCTGCAAGCCAGGCTTTGAAATGTGGGAATACTACTACCACCTTTTGA 7516
2482 GluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLysPro 2501
7517 GAAATGGTCACTGGCTTGGAGGAAAAACCAACATGTAAGCCATTGAGTGCCTCAAAACC 7576
2502 LysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThr 2521
7577 AAGGAGATTTGAATGGCAAAATTCCTTTACCGAGACCTTACACTATGACAGACCGTACC 7636
2522 TyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThr 2541
7637 TACTCTTGCACCGAGGCTTTCGGCTCGAAGTCCCAAGTGCCTTGACCTGTGTAGAGACA 7696
2542 GlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnPro 2561
7697 GGTGATGGGATGTAGATGCCCATCTTGCATGCCATGCCACTGTGATTCGCCCAACACC 7756
2562 IleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleLysSer 2581
7757 ATTGAAATGGTTTTGTAGAAGGTGCAGATTACAGCTATGGTGCATATCATCTACAGT 7816
2582 CysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrp 2601
7817 TGCTTCCCTGGGTTTCAGGTGGCTGGTCACTGCCATGCCAGACCTGTGAAGAGTCAAGATGG 7876
2602 SerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPhe 2621
7877 TCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCCTCCCTCCTCATATAGATTTT 7936
2622 GlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMetMet 2641
7937 GGAGACTGTACTAAACTCAAAAGATGACCGGGATATTTTGAGCAAGAGACGACATGATG 7996
2642 GluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTrp 2661
7997 GAAGTTCCATATGTGACTCTCTCACCTCTCTATCAATTTGGAGCAGTGGCTAAAAACCTGG 8056
2662 GluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetVal 2681
8057 GAAATACAAAGAGGTCTCTCTGCTACACATTCATCAAACTTCTGTATGTGTACCATGTT 8116
2682 SerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGlnGlu 2701
8117 TCATACACCTGTGAATCAGGATATGAATCTCTGGGAACCTCTGTGCTGATCTGCAGAA 8176
2702 AspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThr 2721
8177 GATGGAACTTGGAAATGGCAGTGCACCATCTCTGCAATTCATTAATGTAATGTGACTTCCCTACT 8236
2722 AlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyr 2741
8237 GCTCCTGAAAAATGGCTTTTTCCTTTTACAGAGACTAGCATAGGAGTGGAGTGTGCTGCTAT 8296

QY 2742 SerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArg 2761
DB 8297 AGCTGTAAACCTGGACACATCTAGCAGGCTCTGACTTAAGGCTTTGTCTAGAGAAATAGA 8356
QY 2762 LysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAspVal 2781
DB 8357 AAGTGGAGTGGTCCCTCCACGCTGTGAAGCCATTTCCATGAAAAGCCAAATCCAGTC 8416
QY 2782 MetAsnGlySerIleLysGlySerAsnTrpThrLeuSerThrLeuTrpGluCys 2801
DB 8417 ATGAATGATCCATCAAGAGAACACTACACATACCTGAGCAGCTGTACTATGAGTGT 8476
QY 2802 AspProGlyTrpValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrp 2821
DB 8477 GACCCCGGATATGTCTGAATGGCACTGAGAGGAGAACATGCGCAGATGACAAAACCTGG 8536
QY 2822 AspGluAsnGluProIleCysIleProValAspCysSerProProValSerAlaAsn 2841
DB 8537 GATGAGGATGAGCCCATTTGCAATCTCTGTGGACTGCGAGTTACCCCGAGCTCAGCCCAT 8596
QY 2842 GlyGlnValArgGlyAspGluTrpThrPheGlnLysGluIleGluTrpThrCysAsnGlu 2861
DB 8597 GCCCAGGTGAGAGGAGACGAGTACACATTCGAAAAGAGATTGAATACACTTTGCAATGAA 8656
QY 2862 GlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGly 2881
DB 8657 GGGTTCTTGTCTGAGGGAGCCAGGAGTCCGGTTTGTCTTGGCCAAATGGAGTGGAGTGA 8716
QY 2882 AlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyVal 2901
DB 8717 GCCACTCCCGACTGTGTGCTGTGAGTGTGCAGATGTGCACCCGCCCACTGCCCAATGGGTG 8776
QY 2902 ThrGluGlyLeuAspTrpGlyPheMetLysGluValThrPheHisCysHisGluGlyTrp 2921
DB 8777 ACGGAAGGCTCGACTATGCTTTCATGAAGGAAGTAACATTCACACTGTCCAGGAGGCTAC 8836
QY 2922 IleLeuHisGlyAlaProLysIleuThrCysGlnSerAspGlyAsnTrpAspAlaGluIle 2941
DB 8837 ATCTTGCACGGTGTCTCCAAACCTCACCTGTGAGTCAGATGGCAACTGGGATGCGAGATT 8896
QY 2942 ProLeuLysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsn 2961
DB 8897 CCTCTGTGTAAACCACTCAACTGTGACCTCTGAGACCTCTGAGAGATCTTGCCCATGGTTCCCTAT 8956
QY 2962 GlyPheSerPheIleHisGlyHisIleGlnTrpGlnCysPheProGlyTrpLysLeu 2981
DB 8957 GCTTTTCTCTTATTCATGGGGCCATATACAGTATCAGTGTCTTCTGTTTATAGCTC 9016
QY 2982 HisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSer 3001
DB 9017 CATGGAAATTCATCAAGAAGGTGCCCTCTCCATGGCTCTCGAGTGGCAGCTCACCTTCC 9076
QY 3002 CysLeuProCysArgCysSerThrProValIleGluTrpGlyThrValAsnGlyThrAsp 3021
DB 9077 TGCTGCTCTGCGATGTTCCACACACAGCAATTAATGATATGGAATCTCAATGGACAGAT 9136
QY 3022 PheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeu 3041
DB 9137 TTTTGAAGTGGAAAGGCGAGCCCGATTCAGTGTCTTCAAAAGGCTTCAAGCTCTTAGGACTT 9196
QY 3042 SerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHis 3061
DB 9197 TCTGAATCACCTGTGAAGCCGATGGCAGTGGAGCTCTGGGTTCCTCCACATGTGAAACAC 9256
QY 3062 ThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrp 3081
DB 9257 ACTCTTGTGGTCTCTTCCATGATACCAAAATGCGTTTCATCAGTGGACCACTCTTGG 9316
QY 3082 LysGluAsnValIleThrTrpSerCysArgSerGlyTrpValIleGlnGlySerSerAsp 3101
DB 9317 AAGGAAATGTGATACTTACAGCTGCAGCTGCGATATGTATGTATGCATACAGGAGCTTCAGAT 9376

QY 3102 LeuIleCysThrGluLysGlyValTrpSerGlnProTrpProValCysGluProLeuSer 3121
DB 9377 CTGATTTGTACAGAGAAAGGGGTATGAGCCAGCCTTATCCAGCTGTGAGCCCTTGTTC 9436
QY 3122 CysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTrpGlu 3141
DB 9437 TGTGGGTCCCCACCGTCTGTCCCAATGCAAGTGGCAACTGGAGAGGCACACACATATGAA 9496
QY 3142 SerGluValLysLeuArgCysLeuGluGlyTrpThrMetAspThrAspThrAspThrPhe 3161
DB 9497 AGTGAAGTGAAACTCAGATGTCTGGAAGGTATATCCGATGGATACAGATACAGATACATTC 9556
QY 3162 ThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCys 3181
DB 9557 ACCTGTGAGAAAGATGGTGGTGTCTCCCTGAGAGAAATCTCTGCGAGTCTCTAAATAATGT 9616
QY 3182 ProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArg 3201
DB 9617 CCTCTCCCGAANAACATAACATATACATATGATGGGAGCGATTTTCAGTGTGAATAGG 9676
QY 3202 GlnValSerValSerCysAlaGluGlyTrpThrPheGluGlyValAsnIleSerValCys 3221
DB 9677 CAAGTTCTGTGTGATGTCAGAAAGGTATACCTTTTGGAGGAGTTAACATATCATGATGT 9736
QY 3222 GlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProValSerCys 3241
DB 9737 CAGCTTGATGGAACTTGGGAGCCACCATCTTCCGATGAATCTTTCAGTGTCCAGTTCTTGT 9796
QY 3242 GlyLysProGluSerProGluHisGlyPheValValGlySerLysTrpThrPheGluSer 3261
DB 9797 GGGAAACCTGAAAGTCCAGAACATGATTTGTGTTGGCAGTAAATACACACTTTGAAAGC 9856
QY 3262 ThrIleIleTrpGlnCysGluProGlyTrpGluLeuGluGlyAsnArgGluArgValCys 3281
DB 9857 ACAATTTATTTATCAGTGTGAGCTTGGCTATGAACTAGAGGGGAAACAGGGAGCTGTC 9916
QY 3282 GlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThr 3301
DB 9917 CAGGAGAACAGACAGTGGAGTGGGGGTGGCCATATGCAAGAGACCACTGAGTGTGAACT 9976
QY 3302 ProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnVal 3321
DB 9977 CCACTTGAATTTCTCAATGGAAAGCTGACATTTGAAAAACAGGACGACTGACCCCAAGCTG 10036
QY 3322 ValTrpSerCysAsnArgGlyTrpSerLeuGluGlyProSerGluAlaHisCysThrGlu 3341
DB 10037 GTATATTCCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCTCAGGACCACTGCACAGAA 10096
QY 3342 AsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPhe 3361
DB 10097 AATGGAACTGGAGCCACCCAGTCCCTCTCTGCAAAACCAATTCATGCTGTTCTTTT 10156
QY 3362 ValIleProGluAsnAlaLeuLeuSerGluLysGluPheTrpValAspGlnAsnValSer 3381
DB 10157 GTGATTTCCCGAGATGCTCTGCTGTCTGAAAAGGAGTTTATGTGATCAGAAATGTGTCC 10216
QY 3382 IleLysCysArgGluGlyPheLeuGlnGlyHisGlyIleIleThrCysAsnProAsp 3401
DB 10217 ATCAATATGAGGAAGGTTTTCTGCTGAGGGCCACGGCATCATTTACTGCAACCCCGAC 10276
QY 3402 GluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHis 3421
DB 10277 GAGACTGGACACAGACAGAGCCCAATGTGAAAATAATCTCATGTGTGTCCACCACTCAC 10336
QY 3422 ValGluAsnAlaIleAlaArgGlyValHisTrpGlnTrpGlyAspMetIleThrTrpSer 3441
DB 10337 GTAGAAAATGCAATTCCTGAGGCGTACATTTATCAATATGAGACATGATCACCTACTCA 10396
QY 3442 CysTrpSerGlyTrpMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThr 3461
DB 10397 TGTACAGTGGATACATGTTGGAGGGTTCCTGAGGAGTGTGTTTGTAGAAAATGGAAACA 10456
QY 3462 TrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGlyIle 3481

```

Db      10457 TGGACATCACTCTATTTGGCAGAGCTGTCTGCGATTTCCATGTCAGAAATGGGGCATC 10516
QY      3482 CysGlnArgProAlaCysSerCysProGluGlyTyrMetGlyArgLeuGlyGlu 3501
Db      10517 TGGCAACGGCCAAATCTTTGTTCTGTCTCCAGAGGGCTGGATGGGGCGCTCTGTGAAGAA 10576
QY      3502 ProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCysAsp 3521
Db      10577 CCAATCTGCAATCTTCCCTGTCTGAACGGAGGTGCTGTGTGGCCCTTACCAAGTGTGAC 10636
QY      3522 CysProProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeu 3541
Db      10637 TGGCCGCTGCTGGAGGGGTCTGCTGTCTATACAGCTGTTTGGCCAGTCTCCCTGCTTA 10696
QY      3542 AsnGlyGlyCysValArgProAsnArgCysHisCysLeuSerSerTyrThrGlyHis 3561
Db      10697 AATGGTGGAAATGTGTAAAGACCAACCGATGTCACCTGTCTTTCTTTGGACGGGACAT 10756
QY      3562 AsnCysSerArgCysArgArgThrGlyPhe 3571
Db      10757 AACTGTCTCCAGGAAAGGAGGACTGGGTTT 10786

```

RESULT 8

```

; US-10-107-782-5
; Sequence 5, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Baha,
; APPLICANT: Rekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Shenoy, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glennda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zernusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215

```

```

; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 5
; LENGTH: 11158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (77)..(10786)
US-10-107-782-5

Alignment Scores:
Pred. No.: 0 Length: 11158
Score: 19566.00 Matches: 3503
Percent Similarity: 98.57% Conservative: 16
Best Local Similarity: 98.12% Mismatches: 47
Query Match: 97.96% Indels: 4
DB: 16 Gaps: 2

US-09-977-053-4 (1-3571) x US-10-107-782-5 (1-11158)
QY      4 ArgLeuAlaPheCysCysTyrGlyLeuAlaLeuValSerGlyTyrAlaThrPheGlnGln 23
Db      83 AGAATTTGGCGGGCTTGTGGGGTCTGGCGTCTGTTTGGGCTGGCGACCTTTTCAGCAG 142
QY      24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db      143 ATGTCCCGTTCGGCAATTTTCAGCTTCGGCTCTTCCCGGAGACCGCGCGGGCCCCC 202
QY      44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
Db      203 GGGAGTATCCCGCGCGCGCTCTCTGGGCGACGAGCGGGGAGCAGAGTGGAGCGG 262
QY      64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
Db      263 CTGGCGCGGGCGTTC-----CGGTCGGCTGCTGGGAGCTCAGGAGCGCTGGAG 316
QY      84 LeuValPheLeuValAspSerSerSerValGlyGluValAsnPheArgSerGluLeu 103
Db      317 CTGTCTCTCTGGTGGATGATTCGTCAGCGTGGCGCAAGTCAACTTCCCGCAGCGAGCTC 376
QY      104 MetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db      377 ATGTTCGTCCGCAAGCTGTCTCGACTTCCCGTGGTGGCCACCGCCACCGCGGTGGCC 436
QY      124 IleValThrPheSerSerLysAsnTyrValProArgValAspTyrIleSerThrArg 143
Db      437 ATCTGACCTTCTCGTCCAAAGACTAGTGGTGGCGCGCTGCGATTACATCTCCACCCGC 496
QY      144 ArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIleSerTyrArg 163
Db      497 CGCGCGCGCCAGCACCAAGTGGCGCTGCTCTCCAGAGATCCCTGCCATCTCTTACCGA 556
QY      164 GlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
Db      557 GGTGGCGGCACTTACACCAAGGGCGCTTCCAGCAGCGCGCAAAATCTTCTTTCATGCT 616
QY      184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAsp 203
Db      617 AGAGAAACCTCAACAAAGTTGTATTCTCATCTGATGATATTCATATGGGGGAGAC 676
QY      204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db      677 CCTAGACCAATTCAGCGTCACTGCGAGATTCAGGAGTGGGAGATCTTTCACCTTTGGCATA 736
QY      224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
Db      737 TGGCAAGGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAGGAGCAGCTGT 796
QY      244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArg-----ArgAlaLeuHis 261
Db      797 TACCTGTACACAGTTTTTGAGAAATTTGAGGCTTTTAGTCCGCCCTCTGTGCATATGTTATTT 856
QY      262 GluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeuCys 281

```

Db 857 GTAGATCTACCTTCTGGAGTTTATTCAGATGATATGTCACCTGCTCATATCTTTGT 916
QY 282 AspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGly 301
Db 917 GATGAAGGCAAGAGCTGCTGACCGATGGAAGCTGCAAAATGTGGGAAACACACAGGC 976
QY 302 HisPheGluCysIleCysGluysGlyCysGlyThrGlyGlyGlyLeuGlnThrGluCysThr 321
Db 977 CATTTTAGTGTCTCTGTAAGAGGGGTATACGGGAAAGGCTGACAGTATGACTGCACA 1036
QY 322 AlaCysProSerGlyThrTrpLysProGluGlySerProGlyGlyIleSerSerCysIle 341
Db 1037 GTTTGGCCATCGGGACATACCAACTGGAAGCTCACAGGAGGATCAGAGTGCATT 1096
QY 342 ProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysVal 361
Db 1097 CCATGCTGCTGATAAATCACACTCTCCACCTGGAAGCACATCCCTCGAAGACTGTGTC 1156
QY 362 CysArgGluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeu 381
Db 1157 TCAGAGAGGGATACAGGGGATCTGGCCAGACTGTGAAGTTGTCCACTGCCCTGCCCTG 1216
QY 382 LysProProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAla 401
Db 1217 AAGCCTCCCGAANAATGTTACTTTATCCAAACACTTGCACACCACTTCAATGCAGCC 1276
QY 402 CysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeu 421
Db 1277 TGTGGGCTCCGATGTCACCTCGATTTGATCTTGTGGGAGCAGCATCATCTTATGTCTA 1336
QY 422 ProAsnGlyLeuTrpSerGlySerGluSerTrpCysArgValArgThrCysProHisLeu 441
Db 1337 CCCAATGGTTTGTGGTCCGGTTCAGAGAGCTACTGCAGAGTAAGAACATGTCTCATCTC 1396
QY 442 ArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTrpLysThrThr 461
Db 1397 CGCCAGCCCAACATGCCCCATCAGCTGTTCTACAGGGGAATGTTATATAGACACA 1456
QY 462 CysLeuValAlaCysAspGluGlyTrpArgLeuGluGlySerAspLysLeuThrCysGln 481
Db 1457 TGTTTGGTTGCTGTGATGAAGGTTACAGCTAGAGGCGAGTGATTAAGCTTACTTGTCAA 1516
QY 482 GlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPhe 501
Db 1517 GGAACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGAGCGCCACTGTTCCACCTTT 1576
QY 502 GlnMetProLysAspValIleLeSerProHisAsnCysGlyLysGlnProAlaLysPhe 521
Db 1577 CAGATGCCCAAGATGTCTATCATATCCCCCACTGTGGCAAGCAGCCGCAAAATTT 1636
QY 522 GlyThrIleCysTrpValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMet 541
Db 1637 GGGACGATCTGCTATGTAAGTTGCCCAAGGGTTCATTTTATCTGGAGTCAAGGAATG 1696
QY 542 LeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAsp 561
Db 1697 CTGAGATGTACCACTTCTGGAAATGGAATGTCCGGAGTTCAGGAGCTGTGTAAAGAC 1756
QY 562 ValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnGln 581
Db 1757 GTGGAGCTCTCTCAATCAACTGTCTCTAAGGACATAGAGGCTAAGACTCTGGAACAGCAA 1816
QY 582 AspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysVal 601
Db 1817 GATTCTGCCAATGTTTACCTGGCAGATTCCAACAGCTAAAGACAACTCTGGTGAAGAGGTG 1876
QY 602 SerValHisValHisProAlaPheThrProTrpLeuPheProIleGlyAspValAla 621
Db 1877 TCAGTCCAGTTCATCAGCTTTCACCCCACTTACTTTTTCCCAATGGAGATGTGCT 1936
QY 622 IleValThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLys 641

Db 1937 ATCGTATACACGGCAACTGACCTATCCGGCAACCGAGCCAGCTGCTATTTTCCATATCAAG 1996
QY 642 ValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGlnVal 661
Db 1997 GTTATTGATGCAAGAACCCCTGTCATAGACTGTGTGAGATCTCCACCTCCCGTCCAGGTC 2056
QY 662 SerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAla 681
Db 2057 TCGAGAGAGGTACATCCGACAGCTGGATGAGCTCAGTTCTCAGACCACTCAGGTGCT 2116
QY 682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIle 701
Db 2117 GAATTGCTCATTAACAGAGTATACACAGGAGACTTTTCCCTCAAGGGAGACATATA 2176
QY 702 ValGlnTrpThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
Db 2177 GTACAGTATACAGCCACTGACCCCTCAGGCATTAACAGGAGCATGTGATATCCATATTGTC 2236
QY 722 IleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
Db 2237 ATAAAAGTTCTCCCTGTGNAATTCATTCACACTGTAAATGGGATTTTATATGCACT 2296
QY 742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTrpAspPheThrGlu 761
Db 2297 CCAGATAATACTGCACTCAACTGTACATTAACCTTGTCTGGAGGCTATGATTTTCACAGAA 2356
QY 762 GlySerThrAspLysTrpCysAlaTrpGluAspGlyValTrpLysProThrTrpThr 781
Db 2357 GGGTCTACTGACAAAGTATTATTGTGCTTATGAAGATGGCGTCTGGAACCAACATATACC 2416
QY 782 ThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
Db 2417 ACTGAATGCCAGACTGTGCCAGTAAGCGTTTTCGAAACCCACCGGTTTCAAGTCTCTTGAG 2476
QY 802 MetPheTrpLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAla 821
Db 2477 ATGTTCTACAAAGCAGCTGTTGTGATGACACACAGATCTGATGAAGAAATTTTCTGAAGCA 2536
QY 822 PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
Db 2537 TTTGAGACGACCTCGGAAAAATGGTCCCATCATTTTGTAGTGATGACAGGAGCATTTGAC 2596
QY 842 CysArgLeuGluGluAsnLeuThrLysLysTrpCysLeuGluTrpAsnTrpAspTrpGlu 861
Db 2597 TGCAGACTGGAGGAGAACCTGACCAAAAAATATTGCTAGCATATATATTATGACTATGAA 2656
QY 862 AsnGlyPheAlaIleGlyProGlyTrpGlyAlaAlaAsnArgLeuAspTrpSerTrp 881
Db 2657 AATGGCTTTTGCATTTGCTCCAGGTGGCTGGGCTGCAGCTAATAGGCTGGATTTACTCTTAC 2716
QY 882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
Db 2717 GATGACTTCTTGACACTGTGCAAGAAACAGCCACAGCATCGGCAATGCCAATGCCAGTCTCA 2776
QY 902 ArgIleLysArgSerAlaProLeuSerAspTrpLysIleLysLeuIlePheAsnIleThr 921
Db 2777 CGGATTAAGAGAGTGGCCCATTTATCTGACTATATAAATTAAGTTAATTTTAAACATCACA 2836
QY 922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGluArg 941
Db 2837 GCTAGTGTGTCATTACCCGATGAAGAAATGATACCTTGAATGGGAAAAATCAGCAACGA 2896
QY 942 LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961
Db 2897 CTCCTTCAGACATTTGAAACTATCAAAATTAATGAAAGGACTCTCAACAAAGACCCC 2956
QY 962 MetTrpSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
Db 2957 ATGTATTTCTTTTCCAGCTTGCATCAGAAATATCTTATAGCCGACAGCAATTTATTAGAAACA 3016
QY 982 LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
Db 3017 AAAAAGGCTTCCCTTTCTGCAGACAGGCTCAGTGTCTGAGAGGGCGGTATGTGTGTCAAT 3076

Qy	1002	CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGly	1021
Db	3077	TGCCCTTTGGGAACCTATTATAATCTGGAAACATTTACCTGTGAAGCTGCCGATCGGA	3136
Qy	1022	SerTyrGlnAspGluGluGlnLeuGluCysAlaValCysProSerGlyMetTyrThr	1041
Db	3137	TCCCTATCAAGATGAAGAGGCACTTGAATGCAAGCTTTGCCCTCTGGGATGTACACG	3196
Qy	1042	GluTyrIleHisSerArgAsnIleSerAspCysAlaGlnCysIysGlnGlyThrTyr	1061
Db	3197	GAATATATTCATTCCAGAAACATCTCGATTGTAAAGCTCAGTGTAAACAGGACCACTAC	3256
Qy	1062	SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProIysPhe	1081
Db	3257	TCATACAGTGGACTTGAAGCTTGTGAATCTGTGTCACCTGGGCACTTATCAGCCAAATTT	3316
Qy	1082	GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValIysArgGlyAlaVal	1101
Db	3317	GGTTCCCGAGCTGCCCTCTCGTGTCCAGAAAACCTCAACTGTGAAAAGAGGAGCGTG	3376
Qy	1102	AsnIleSerAlaCysGlyValProCysProGluGlyIysPheSerArgSerGlyLeuMet	1121
Db	3377	AACATTTCTGCATGTGGAGTTCTTGTCAGAGGAATAATCTCGCGTTCTGGTTAATG	3436
Qy	1122	ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyIysAlaPheCysLeu	1141
Db	3437	CCCTGTCAACCATGCTCTCGTGACTATTACCAACCTAATCAGGGAAGGCTTCTGCCTG	3496
Qy	1142	AlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSer	1161
Db	3497	GCCTGTCCCTTTTATGGAACTTACCCTCAATCTCGTGTCCAGATCCATCAGAGAAATGTTC	3556
Qy	1162	SerPheSerThrPheSerAlaAlaGluSerValProProAlaSerLeuGly	1181
Db	3557	AGTTTTAGTTCACTTTCTCAGCGGAGAGGAAGTGTGTGCCCCCTGCTCTCTTGA	3616
Qy	1182	HisIleIysIysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro	1201
Db	3617	CATATTAAAGAGGCGATGAATCAGCAGTCAGGCAAGTCATGAATGCTTCTTTAACCTCT	3676
Qy	1202	CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu	1221
Db	3677	TGCCACAATAGTGAACCTGCGACACACTTGGGCGTGTATTANGTTTGTCTCTGTGCCATT	3736
Qy	1222	GlyTyrThrGlyLeuIysCysGluThrAspIleAspGluCysSerProLeuProCysLeu	1241
Db	3737	GGATATACAGGTTTAAAGTGTGAACACAGACATCGATGAGTCAGCCCACTGCCCTTGCCTC	3796
Qy	1242	AsnAsnGlyValCysIysAspLeuValGlyluPheIleCysGluCysProSerGlyTyr	1261
Db	3797	AACAAATGGAGTTCTTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGCCCATCAGGTTAC	3856
Qy	1262	ThrGlyGlnArgCysGluGluAsnIleAsnGlnCysSerSerSerProCysLeuAsnIys	1281
Db	3857	ACAGTAAGCACTGTGAATTTGAACATCAATGAATGTCTGAGTCAATCTAGAAATCAG	3916
Qy	1282	GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValIysGlyPheValGly	1301
Db	3917	GCCACCTGTGTGGATGAATTAATAATTCATACAGTTGTAAATGTCTCAGCCAGGAATTTTCAGGC	3976
Qy	1302	LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal	1321
Db	3977	AAAAGGTGTGAAACAGGTATGTATCAACTCAGTGTATTATTAATCACTTATTAATGCGATC	4036
Qy	1322	CysGluAspGlnValGlyGlyPheLeuCysIysCysProGlyPheLeuGlyThrArg	1341
Db	4037	TGTGAAGACCAGGTTGGGGGATCTTGTGTCAAAATGCCCAACCTGGATTTTTTGGGTACCCGA	4096
Qy	1342	CysGlyIysAsnValAspGluCysLeuSerGlnProCysIysAsnGlyAlaThrCysIys	1361
Db	4097	TGTGGAAGAAACGTGATGATGTCTCAGTCAGCCATGCAAAAATGAGTCACTCTGTAAA	4156

QY	1362	AspGlyAlaIaIaSerPheArGysLeuCyCyAlaAlaGlyPheThrGlySerHisCysGlu	1381
DB	4157	GACGGTGCCTAAATAGCTTTACGGTCTGCTGTGTGCAGCTGGCTTCAACAGGATCACACTGTGAA	4216
QY	1382	LeuAsnIleAsnGluCyGlnSerAsnProCysArGAsnGlnAlaThrCysValAspGlu	1401
DB	4217	TTGAACATCAATGAATGTCAGTCTAATCTCAATGAGAAATCAGGCCACCTGTGTGGATGAA	4276
QY	1402	LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu	1421
DB	4277	TTAAATTCATACAGTCTTAAATGTTCAGCCAGGATTTTCAGGCCAAAAGGTGTGTGAACACAA	4336
QY	1422	GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu	1441
DB	4337	CAGTCTACAGGCTTTAAACCTGGATTTTTTGAAGTTTCTGGSCATCTATGGATATATGCTAGCTA	4396
QY	1442	AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAsp	1461
DB	4397	GATGGCATGCTCCCATCTCTCCATGCTCTAACTGTACCTTTCTGGATGAATATCTCTGCAC	4456
QY	1462	AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeu	1481
DB	4457	GACATGAACATATGGAAACACCAATCTCTCATGAGTGTATTAACGCGACGACAAATACCTTG	4516
QY	1482	LeuLeuThrAspTyrAsnGlyTyrPValLeuTyrValAsnGlyArgGluLysIleThrAsn	1501
DB	4517	CTCTCTGACTGATTAATACGGGTGGTTCCTTATGTGAAATGGCAGGGAAGATAACAAC	4576
QY	1502	CysProSerValAsnAspGlyArgTyrPheHisIleAlaIleThrTyrThrSerAlaAsn	1521
DB	4577	TGTCTCTCGTGGTAATGAGATGGCAGATGGCATCATATGTCATATCACTTGGACAGTACTGGT	4636
QY	1522	GlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerVal	1541
DB	4637	GGCAAGCCATATCTGTGTGGCGGTGCATATAGTTCTCTGGGCAAGACAGCAAAAAAGGA	4756
QY	1542	GlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluAspLysLysGly	1561
DB	4697	GGCAAGCCATATCTGTGTGGCGGTGCATATAGTTCTCTGGGCAAGACAGCAAAAAAGGA	4756
QY	1562	GluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAsp	1581
DB	4757	GAGGGGTTCACCCGGCTGAGTCTTTTGTGGGCTCCATAGCCACAGCTCAACCTCTGGGAC	4816
QY	1582	TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer	1601
DB	4817	TATGTCTGTCTCCACAGCAGGTGAAGTCACTGGTACTCTCTGCCACAGGAACCTCAGT	4876
QY	1602	LysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLysIle	1621
DB	4877	AAAGCAACCTGTGTAGCATGCGCTGATTTCTTGTTCAGGAATGTGGGGAAAGTGAAGATC	4936
QY	1622	AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeu	1641
DB	4937	GAITCTAAGAGCATATTTTGTTCATGTTCGATTCGCCACGCTTGGAGGGTCAGTGCCTCATCTG	4996
QY	1642	ArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly	1661
DB	4997	AGAACTGCATCTGAAGATTTTAAACACAGTTCACAACTCATCTGTCTGTGAACCCAGGC	5056
QY	1662	PheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnIleTyrThrGlnPro	1681
DB	5057	TTCCAGCTGGTGGGAAACCCCTGTGTGCAGTACTGTCTGAATCAAGACACAGTGGACACAAACCA	5116
QY	1682	LeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHis	1701
DB	5117	CTCCCCCATGTGTGAACGCATTTGCGCTGTGGGGTGGCACCTCTCTTTGGAGAAATGGCTCCAT	5176
QY	1702	SerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyr	1721
DB	5177	TCAGCCCATGACTCTATGCTGGCGGACAGTAACTTACCTGCTGGTGCAACAATGGCTACTAT	5236
QY	1722	LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerPro	1741

2462 AlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeuCysGly 2481
DB GCTCTCTATACCTGCAAGCCAGCTTTGAATTGGTGGGAATACTACCAACCTTTTGTGA 7516
2482 GluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLysPro 2501
DB GAATAATGGTCACTGGCTTGGAGGAACCAACACATGTAAGCCATTGAGTGCCTGNAACCC 7576
2502 LysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThr 2521
DB AAGGAGATTTCGAATGGCAAAATTCCTTACACGACCTACACTATGACAGACCGTTACC 7636
2522 TyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThr 2541
DB TACTCTTGCACCGAGCTTTTGGCTCGAAGGTCCCGAGTCCCTGTGACCTGTGTAGAGACA 7696
2542 GlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnPro 2561
DB GGTGATTGGGATGATAGTCCCATCTTGCATGTCCTCCACTGTGATTTCCCCACAAACC 7756
2562 IleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleLeuTyrSer 2581
DB ATTTGAAATGGTTTGTAGAAAGTGCAGATTACAGCTATGTGTGCCATAATCTACAGT 7816
2582 CysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrp 2601
DB TGTCTCCCTGGGTTTCAGTGGCTGTGTGATGCTCATGCCATGACAGCTGTGAAGAGTCAGATGG 7876
2602 SerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPhe 2621
DB TCAAGTTCCATCCCAACATGATGCCAATAGACTGTGGCTCCCTCCTCATATAGATTTT 7936
2622 GlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMetMet 2641
DB CGAGACTGTACTAACTCAAGATGACAGGGATATTTTGGAGCAAGACGACATGATG 7996
2642 GluValProTyrValThrProHisProProTyrHisLeuGlyValAlaValAlaValThrTrp 2661
DB GAAGTTCCATATGTGACTCCTCACCTCTCTATATCTTGGAGCAGTGGCTAAACCTGG 8056
2662 GluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetVal 2681
DB GAATAATACAAAGAGTCTCTCTGCTACACATTCATCAAACTTTCTGTATGTACCATGGTT 8116
2682 SerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGlnGlu 2701
DB TCATACACCTGTAATCCAGGATATGAATCTCTGGGGAACCTGTGTGCTGATCTGCCAGGAA 8176
2702 AspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThr 2721
DB GATGGAACCTTGGAAATGGCAGTGCACCATCTCGCATTTCAATGGAATGTGACTTGCCTACT 8236
2722 AlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyr 2741
DB GCTCTCTGAAATGGCTTTTGGTTTACAGAGCTAGCATGGGAAGTCTGTGTCAGTAT 8296
2742 SerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArg 2761
DB AGCTGTAAACCTGGACACATCTAGCAGCTCTGACTTAAGGCTTTGTCTAGAGATAGA 8356
2762 LysTrpSerGlyValAsnSerProArgCysGluAlaIleSerCysLysLysProAsnProVal 2781
DB AAGTGGAGTGGTGGCTCCCGCTGTGAAGCCATTTCTATGCAAAAGCCAAATCCAGTC 8416
2782 MetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCys 2801
DB ATGATATGGATCCATCAAAGAGAGCAACTACACATACCTGAGCAGCTGTGTACTATGAGTCT 8476
2802 AspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrp 2821
DB GACCCCGGATATGTCTGAATGGCATGTAGAGAGAGAAATGTCAGGATGACAAAACTGG 8536

2822 AspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAlaAsn 2841
DB GATGAGATGAGCCCATTTGCAATCTCTGTGCACTGCAGATTCAACCCAGTCTCAGCCAT 8596
2842 GlyGlnValArgGlyAspGluTyrThrPheGlnGlyLeuIleGluTyrThrCysAsnGlu 2861
DB GGCAGGTGAGAGAGAGAGTACATTCACAAAAAGAGATTGAATACACTTGCATGCAAA 8656
2862 GlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGly 2881
DB GGGTTCTTGTCTGAGGAGCCAGAGTGGGTTTGTCTGCCAATGGAAGTTGGAGTGA 8716
2882 AlaThrProAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGlyVal 2901
DB GCCACTCCCGACTGTGTGCTGTCCAGATGTGCCACCCGCCACAACTGGCCAAATGGGGT 8776
2902 ThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyr 2921
DB ACGGAAGGCTGACTATGTCTTCATGAAGAAAGTAACATTCACCTGTCCAGAGGGCTAC 8836
2922 IleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIle 2941
DB ATCTTCACAGGTGCTCCAAACCTCACCTGTCACTCAGATGGCACTGGGATGCAAGATT 8896
2942 ProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsn 2961
DB OCTCTCTGTAAACCACTCACTGTGGACCTCTCTGAAGATCTTGGCCATGGTTTCCCTAAT 8956
2962 GlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeu 2981
DB GGTTTTCTTTTATTCATGGGGCCATATACAGTATCAGTGTCTTCTGTGTATATAAGCTC 9016
2982 HisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSer 3001
DB CATGGAAATTCATCAAGAGGTGCTCTCCAAATGGCTCTCTGGAGTGGCAGCTCACCTTCC 9076
3002 CysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAsp 3021
DB TGCCTGCTTGCAGATGTTCACACCAAGTANTTGAATGTGAACCTGTCAATGGGACAGAT 9136
3022 PheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeu 3041
DB TTTGACTGTGGAAGGAGCCCGGATTCAGTGTCTTCAAAGGCTTCAAAGCTCTCAGGACTT 9196
3042 SerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHis 3061
DB TCTGAATACCTGTGAAGCCGATGGCAGTGGAGCTCTGGGTTCCTCCCTGTGAACAC 9256
3062 ThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrp 3081
DB ACTTCTTGTGTCTCTTCCATGATACCAATGGCTTCATCAGTGAGACCGAGCTCTTGG 9316
3082 LysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAsp 3101
DB AAGGAAATGTGATAACTTACAGCTGCAGCTCTGGGATATGTCTATACAAGGACAGTTCAGT 9376
3102 LeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSer 3121
DB CTGATTTGTACAGAAAGGGGTATGAGCCGCTTATCAGTCTGTGAGCCCTTGTTC 9436
3122 CysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGlu 3141
DB TGTGGTCCCGCTCTGTGGCCATGCACTGGCACTGGAGAGGGCACACACCTATGAA 9496
3142 SerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPhe 3161
DB AGTGAAGTGAACCTCAGATGTCTGGAAGGTTATACGATGGATACAGATACAGATACATTC 9556
3162 ThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCys 3181
DB ACTGTGCAGAAAGATGGTGGTGGTTCCCTTGAGAGAAATCTCTCGAGCTCTCAAAAAATGT 9616
3182 ProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArg 3201

Db 2870 TGTCTACGGATTAAAGAACTGTCCCATTTGTCTGACCCCAAAATTCAGCTAATTTTTTAAC 2929
Qy 920 ILeThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGln 939
Db 2930 ATCACAGCTAGCGTGCCACTCCACAGAGGAAAGAAACGATACCTTTGAAATTGAGAAATCAG 2989
Qy 940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959
Db 2990 CAGCGACTATTAGACATTTGAAACAATCACCAATCGCTGAAAGCACTTGTGATATA 3049
Qy 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
Db 3050 GAGCCCATGATTCTTCCAGCTCGCTCGGAAACAGTGTGTGTCGACAGCAATTTCCCTC 3109
Qy 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
Db 3110 GAAACAGAAAAGGCTTTTCTCTCTCGACAGCAGGCTCTGTCTGAGGGGGGCGCATGTGT 3169
Qy 1000 ValAsnCysProLeuGlyThrTyrTrpAsnLeuGluHlePheThrCysGluSerCysArg 1019
Db 3170 GTCAACTGCCCCCTGGGAACCTTCTACTCTCTGGAGCATTCACCTGTGAAAGCTGCCCC 3229
Qy 1020 IleGlySerTyrGlnAspGluGlnGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
Db 3230 ATGGATCTTACCAAGATGAAGAGGCGAGCTGGAATGCAAGCTCTGTCCCCCAAGGACT 3289
Qy 1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
Db 3290 CACGGCGAATATCTCCATTCAGAAAGCGTCTCTGAATGCAAAAGCTCAGTGTGAAGCAAGCG 3349
Qy 1060 ThrTyrSerTyrSerGlyLeuGlnThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
Db 3350 ACCTACTCTTCAGTGGGCTGGAGACTCGGAATCGTGTCCCTGGGTACTTATCAACCG 3409
Qy 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
Db 3410 GAATTTGGATCCGGAGCTGCTCTTATGCCAGAAACCAACCAAGCGTGAAGAGGA 3469
Qy 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
Db 3470 GCGGTGGACATCTCTGCTGTGGAGTGGCTTCCACAGTAGAGAGAAATTCCTCCGTTCTGGG 3529
Qy 1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
Db 3530 CTAAACACCTGTGTACCTTGTCCCTCGAGACTATTACCAACCCCAATGACGGAGTCTTTC 3589
Qy 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
Db 3590 TGCTCAAGTTTAGCTCTACTTCTCAGCAGCAGAGAAAGCATAGTGCCTCGTGGCC 3649
Qy 1160 CysSerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSer 1179
Db 3650 TGCTCAAGTTTAGCTCTACTTCTCAGCAGCAGAGAAAGCATAGTGCCTCGTGGCC 3709
Qy 1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe 1199
Db 3710 CTGGACATCTCCAGAACATGACGAGTACGAGTCAAGTCTTTCACCAAGTCTCTTA 3769
Qy 1200 AsnProCysHisIleAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCys 1219
Db 3770 AACCCCTGCCAACACAGTGGAACTTCCAACTGCAACAGCTTGGGGGTGTTATGTCTGTCTGC 3829
Qy 1220 ProLeuGlyTyrThrGlyLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
Db 3830 CCACCTGGATACACAGGCTTAAAGTGTGAACAGATATTGATGAATGCAAGCTCTCTGGCT 3889
Qy 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
Db 3890 TGCTCAATGGTGAATTTGTAGAGACCAAGTTGGGGGATTCAGTGTGCAGATTTCAATTG 3949
Qy 1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeu 1279
Db 3950 GGCTATTTCAGTCAATATGTGAAGAAATATAAATGATGTGTATCTCCAGGCCCTCTCTTA 4009

Qy 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
Db 4010 AATAAAGGAACCTGCACTGACCGCTTGGCAAGCTACCGCTGTACCTGTGTGAAGGATAC 4069
Qy 1300 ValGlyLeuHisCysGlyThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
Db 4070 ATGGGTGTGCACCTGTGMAACACAGCTCAATGAATGCCAGTCAAGCCCTCTTAAACAC 4129
Qy 1320 AlaValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGly 1339
Db 4130 GCAGTTTGTAAAGACCAAGTTGGGGGTCTCTGTGCAATATGCCACCCCGATTTTGGGT 4189
Qy 1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr 1359
Db 4190 ACTCGGTGTGAAAAAATGTGATGAGTGTCTCAGTCAACCATGCCCAATATGANGCCT 4249
Qy 1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
Db 4250 TGTAAAGGATGTGTCACACAGCTTCAGTGTCAATGTCCAGCAGGCTTTCACAGGACAC 4309
Qy 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
Db 4310 TGTGAACGAAACATCAACGAGTGTGAGTCCAAACCCGTGTAGGAACCCAGGCCACCTGTG 4369
Qy 1400 AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db 4370 GATGAACATACTCATACAGTTGTAAATGTACAGCAGGATTTTCAGGCCACAGGTGTGAG 4429
Qy 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
Db 4430 ACAGAACACGCTTCGGGTTTTAACCTGGATTTTGAAGTTTCTGGCATCTACGGGTACGTC 4489
Qy 1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSer 1459
Db 4490 CTGCTAGATGGAGTGTGCAACCCCTCCATCGCTAAACCTGCGCATCTTCGATGAAATCC 4549
Qy 1460 SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn 1479
Db 4550 TCTGATGTCTCACTACACTACGGAGCGCCATCTCTATGCACTTGAGGATGCAAGACAC 4609
Qy 1480 ThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGlyLysIle 1499
Db 4610 ACCTTCTCTCACTGATTACAAAGCGCTGGTCTCTTATGTGAATGCAAGAAAGATC 4669
Qy 1500 ThrAsnCysProSerValAsnAspGlyArgTyrHisHisIleAlaIleThrTyrThrSer 1519
Db 4670 ACCAACTGCCCCCTCCGTAATGATGTCATTTGGCATCATATTGCAATTCACATGGACA 4729
Qy 1520 AlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeu 1539
Db 4730 ATTGGTGGAGCTTGGAGGCTCTATATAGATGGGAATTTATCTGACCGTGTGTACTGGCTC 4789
Qy 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluInAspLys 1559
Db 4790 TCCATTGGCAAGGCCATCATCTGTGCGGTGTCATTAGTTCTTGGGCAAGACAGACAA 4849
Qy 1560 LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579
Db 4850 AAGGAGAGGGGTTCACACCGGCTGAGTCTCTTTGTGGGTCTCCATAAGCCAGCTCAACCTC 4909
Qy 1580 TrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4910 TGGGACTATGTCTGTCTCAACAGCAGGTGAAGTGTCTGGCAGCTCTCTGCCAGAGAA 4969
Qy 1600 LeuSerLysGlyAsnValLeuAlaIleProAspPheLeuSerGlyIleValGlyLysVal 1619
Db 4970 CTGAGTCCGGGAAACGTTGTAGCATGGCCCGATTTCTGTCTGGGAATTCACGGGAGGTG 5029
Qy 1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro 1639
Db 5030 AAGGTTGATTCAGACGACATGTTCTCTGTGATGTCCGCTCTTTAGANGGATCCGTCCT 5089

1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
1659 CACCTGAGACCTGCATCAGGAAATCGAAGCCAGGCTCCAAAGTCAGTCTGTCTGTGAT 5149
1660 ProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThr 1679
5150 CCGGGCTCCAGATGGTGGGAATCTGTGTCAGTATTGTCTGAACCAAGCGCAGTGGACA 5209
1680 GlnProLeuProHisCysGluAlaGlnIleSerCysGlyValProProProLeuGluAsnGly 1699
5210 CAACCACTCCCGCACTGTGAACGCAATTCGCTGTGGGCTGCCCTCCCGCTTTGGAGAAATGGC 5269
1700 PheHisSerAlaAspAspPheTyrAlaGlySerThrValTyrTyrGlnCysAsnAsnGly 1719
5270 TTCTACTCAGCCGAGGACTTCCATGCGGGCAGCACCGGTGACCTTTCAGTGCCACCATGGC 5329
1720 TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyVal 1739
5330 TACTACTCTGCTGGTGAATTCCTCCGAATGTTCTGCRACAGAACCGGAGCTGGAACGGCAAT 5389
1740 SerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAla 1759
5390 TCACCACTCTGCTCGATGTGATGATGTCAGTCGGCTCGGACTGTATGTAGACACGCC 5449
1760 SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp 1779
5450 TCTTGCTCTGACCAACCAACGATCTCTACGATGCTCTGTAAACCACCATACACGGAGAT 5509
1780 GlyLysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHis 1799
5510 GGGAAAACTGTGCAGAACCTGTGAATAATGTAAAGGCTCCAGNAATCCAGAAATGCGCCG 5569
1800 SerSerGlyGluIleTyrThrValGlyAlaValThrPheSerCysGlnGluGlyTyr 1819
5570 TCTTCTGGCGAGATTTACACCGTGGGTACTGTCAGTGCACATTTTCTGTGACGAAGGGCAC 5629
1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGlnTyrAsnHisLeuIle 1839
5630 GAGCTGGTGGAGTGGACCATCAGCTGTGTGGAGCTGGCGAGTGGGATCGCTCAGG 5689
1840 ProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGlu 1859
5690 CCGTCTGTGAAGCCATTTCTGTGTGTGTCCTCCACCTGTTCTGAAATGCTGTGTGAC 5749
1860 GluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeu 1879
5750 GGGTCGGCAATTCACATATGCGAGTAAAGGTGGTGTACAGGTGTGATAAAGGATATACTTTG 5809
1880 AlaGlyAspLysSerCysLeuAlaAsnSerSerTyrSerHisSerProProVal 1899
5810 TCTGGGATGAAGAGTCAGATGCTTGTAGTGTCTCTGGAGTCAATTCCTCTCTCTGTG 5869
1900 CysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSer 1919
5870 TGGGGCTAGTGAAGTGTTCCTCCAGCTGAGGACATAAATAACGCAAAATACATCTTAAGT 5929
1920 GlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGly 1939
5930 GGGCTCACCTACCTTCTATTGTCATGCTACTCTGTGAGAACCGATACAGTTTACAGGGC 5989
1940 ProSerIleIleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHis 1959
5990 CCATCTCTCTTGAATGACAGCTTCCGGCAGCTGGGACAGCGCCACCTAGCTGTCAA 6049
1960 LeuValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPhe 1979
6050 CTGTCTCTCGCGAGAGCTCCCAATCGTCARAGATGCTGTATCATCTGGGACCACTTC 6109
1980 ThrPheArgAsnThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLeuAsp 1999
6110 ACTTTTGGGAACACAGTTGTTACATGACAAAGAGGGCTACACCTTGTGGGCGCTGAC 6169
2000 ThrIleGluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaVal 2019

6170 ACCATCATATGCCAGGCCAACGCAATTCAGTAATCAACCACTGCTGGCTGTC 6229
2020 SerCysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhe 2039
6230 TCTGTGACGAGCCGCCCAATGTGGACCAACGCTCTCCAGAGACTGCTCAGAGCTCTTT 6289
2040 GlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeu 2059
6290 GGAGACACCGCTTTTACTACTGTGGATGGCTACAGCTTGGCTGATTAATTCACAGCTC 6349
2060 LeuCysAsnAlaGlnGlyLysTyrValProProGluGlyGlnAspMetProArgCysIle 2079
6350 ATCTGCAATGCCAGGGGAATCTGGTTTCCCGCGGCCAGGCTGTGCGCGCTGCATA 6409
2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLys 2099
6410 GCTCACTTCTGTGAAACACCCCATCTGTTCCTACAGCATCTTGGAAATCTGTGAGCAA 6469
2100 AlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsn 2119
6470 GCAAGTTTTCAGCTGCTCGGTAGTGAGCTTCAAGTGCATGGAGGTTTGTGCTGAAC 6529
2120 ThrSerAlaLysIleGluCysMetArgGlyGlyGlnTyrAsnProSerProMetSerIle 2139
6530 ACCTCAGCGAGATTCATGCTGAGAGGTGGAGAGTGGAGCCCTTCTCCCTCTCGGTC 6589
2140 GlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGly 2159
6590 CAGTGCATCCGCTGCGATGCGGAGAGCTTCCAAAGCATCGCAATGCTACCCGAGTGG 6649
2160 SerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLys 2179
6650 ACAACTACAGTTTGGGCGCTGGTGCTCAGCTGCCACAGAGGATTTCTATATCAG 6709
2180 GlyLysLysSerThrCysGluAlaThrGlyGlnTyrSerSerProIleProThrCys 2199
6710 GGGGAGAGAGAGACGCTGTGAGGCCACAGGACAGTGGAGTAAACCCACGCCCTGTC 6769
2200 HisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThr 2219
6770 CATCTCTGTCTGTAAACGAGCCACTTAAGTTTGAGAACGGCTTCTCTGGAGCACACCAT 6829
2220 GlyArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGly 2239
6830 GGCAGACCTTTGAGAGCGAAGCAGGTTCCAGTGCACACCCAGCTATAAGGCGACCGGA 6889
2240 SerProValPheValCysGlnAlaAsnArgHisTyrHisSerGluSerProLeuMetCys 2259
6890 AGTCTGTGTTGTTTGGCAAGCCAAATCGCCACTGGCACAGCGCCCTCTGTCTGTC 6949
2260 ValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsn 2279
6950 ACCCTCTCACTGTGGGAAACCCCTCCCAATTCAGAAATGGCTTTTGAAGGAGAAAGC 7009
2280 PheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAsp 2299
7010 TTTGAAGTGGGTCCAGGTTTCAAGTTTGTCTGTAAAGAGGATATGAGCTCGTTGGTAT 7069
2300 SerSerTyrThrCysGlnLysSerGlyLysTyrAsnLysLysSerAsnProLysCysMet 2319
7070 AATCTTGGACTTGGCAAAATCTGGCAATGAGTAAAGAACCAAGCCCGAGTGTGTC 7129
2320 ProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThr 2339
7130 CCCACCAAGTGTGACAGAGCTCTCTCTTAGAAAAACAGCTCGTATTGAAGGAATTAGCT 7189
2340 ThrGluValGlyValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSer 2359
7190 TCCGAGTGGAGTGAAGACCATTTCTGTAAAGAGGGCATGCTTGGAGGCCCTCT 7249
2360 ValLeuLysCysLeuProSerGlnGlnTyrAsnAspSerPheProValCysIleVal 2379

Db 7250 GTCCTGAAGTCTGCCATCCGGCAATGGATGGTTCCTTTCTTATTTGTAAGATGGTC 7309
Qy 2380 LeuCysThrProProLeuLeuSerPheGlyValProLeuProSerSerAlaLeuHis 2399
Db 7310 CTTTGTCCCTCGCCCTCCCTTGAATCCCTTGGCGTCCCTGCTCTTCCGGTCTCTTCAT 7369
Qy 2400 PheGlySerThrValLysThrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThr 2419
Db 7370 TTTGGCAGTACTGTCAAGTATCTGTGTGTGACGGGTCTTTCTTAAGAGGCGATCCAAACC 7429
Qy 2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
Db 7430 ATCCCTCTCCAGGCTGTATGACACCTGGAGTCTCCATTTGCCCGAATGGCTTCGGTAGAA 7489
Qy 2440 CysProGlnProGluGluLeuPheProAsnGlyLeuLeuAspValGlnGlyLeuAlaTrpLeu 2459
Db 7490 TGTCCCAACCTGAGGAGATCTCAACGATATCATCCAGCTACAAAGGGCTTGCTTATCTC 7549
Qy 2460 SerThrAlaLeuThrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeu 2479
Db 7550 AGCACACGCTCTACACCTGCAAGCCAGCTTTGAGTTAGTGGGCAATGCTACCCACCTC 7609
Qy 2480 CysGlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaLeuLeuGluCysLeu 2499
Db 7610 TGTGGGGAATATGGCCAGTGGCTCGAGGAAACCAATGTGCAACCCATTGATGCCCA 7669
Qy 2500 LysProLysGluLeuAsnGlyLysPheSerThrThrAspLeuHisTrpGlyGlnThr 2519
Db 7670 GAGCCCAAGGAGATTTTAAATGGCCAACTCTCTCCGTGAGCTTTTCAGTATGACAAACC 7729
Qy 2520 ValThrTrpSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
Db 7730 ATCATATCTTTTGTGACCGGGGCTTCGGCTCGAAGGTCCCAATTCCTGACCTGTTTA 7789
Qy 2540 GluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaLeuHisCysAspSerPro 2559
Db 7790 GAGACAGTGTGATGGATATGGATCCCTCTCTGTGATGCCATCCACTGCAAGTACCCA 7849
Qy 2560 GlnProLeuGluAsnGlyPheValGluGlyAlaAspThrSerThrGlyAlaLeuLeuLeu 2579
Db 7850 CAGCCCATTTGAAATGGTTTGTAGAGGTGGGATTAACAGATACGGTGGCCATGATCATC 7909
Qy 2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
Db 7910 TATAGTCTCTCCCTGGGTTCAGTGTCTGTGTCATGCCATGACACCTGTGGAAGTGG 7969
Qy 2600 GlyTrpSerSerIleProThrCysMetProLeuAspCysGlyLeuProProHisIle 2619
Db 7970 GATGTCTCAAGCTCAGCCCAACCTGTGTACCCATAGACTGGGTCTCCCTCTCCACATA 8029
Qy 2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyThrPheGluGlnGluAspAsp 2639
Db 8030 GACTTTGTGTGCTGTACTAAAGTCAGAGATGCCAGGGACATTTTGTATCAAGAGATGAC 8089
Qy 2640 MetMetGluValProThrValThrProHisProProThrHisLeuGlyValAlaLys 2659
Db 8090 ATGATGGAAGTCCCATATCTGGCT-----CACCTCAACATTTTGAAGCAACAGCTAAG 8143
Qy 2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTrpGlyThr 2679
Db 8144 GCCTTGGAATATCAAGAGTGGCTGCTCCATGATGCCATCCACTTCTCTATGGCAGC 8203
Qy 2680 MetValSerThrCysAsnProGlyThrGluLeuLeuGlyAsnProValLeuLeuCys 2699
Db 8204 ATGGTTCTCTACAGCTCGAGCTGGTTATGAACTGCTGGGAATCCCTGTGCTGATCTGC 8263
Qy 2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
Db 8264 CAGGAAGATGTGTGTAATGGTATCCGACCTCTTGCATTTCCATTTGATGATGATTTG 8323
Qy 2720 ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
Db 8324 CTTGTGTCTCCCGAAATGGCTTTTATCATTTTACATTTTACACAGACGACTATGGGCAGTGTGCA 8383

Qy 2740 GlnTrpSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
Db 8384 CAATATAGCTGCAAGCCCGGGCACATTTAGAAAGGCTCCACCTTAGAGACTGTGTCTGCAG 8443
Qy 2760 AsnArgLysTrpSerGlyValAserProArgCysGluAlaIleSerCysLysLysProAsn 2779
Db 8444 ATATAGCAGTGGGTGGCAGCTGTTCCAGCTGTGAAGCCATCTCATGAGTAAAGCCAAAC 8503
Qy 2780 ProValMetAsnGlySerIleLysGlySerAsnTrpThrTrpLeuSerThrLeuTrpYr 2799
Db 8504 CCACCTCGAATGGATCCATCAAGAGATGACTACTCTCTACCTGGGTGTGTATACTAC 8563
Qy 2800 GluCysAspProGlyTrpValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLys 2819
Db 8564 GAGTGTGACTCTGGCTATATTTCTCAATGGCTCTAAGAGAGGACATGCCCAAGAAATAGA 8623
Qy 2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
Db 8624 GATTGGATGGCATGAGCCCATGTGTATCTCTGTAGACTGTGGCTCACCCCGAGTCCCC 8683
Qy 2840 AlaAsnGlyGlnValLargGlyAspGluTrpPheGlnLysGluLeuGluTrpYrCys 2859
Db 8684 ACCAATGCCGAGTCAAGGAGAGATACATTTCCAAAGAGATTTATACATCTCTTGC 8743
Qy 2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
Db 8744 CGTGAAGGTTTCATCTCGAAGGAGCCAGAGTGTGTATCTGTCTTACCAATGGAGTTGG 8803
Qy 2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProGlnLeuAlaAsn 2899
Db 8804 AGTGTGTGCATCTCCAGCTGTGATGTGTAGATGTCTGCCCCCACCACAGGTGCCAAT 8863
Qy 2900 GlyValThrGluGlyLeuAspTrpGlyPheMetLysGluValThrPheHisCysHisGlu 2919
Db 8864 GGGGTGGCAGTGGCTAGACTATGGGTTCAAGAGAGAGTAGCGTTCACATGTCTTAGAG 8923
Qy 2920 GlyTrpIleLeuHisGlyValaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
Db 8924 GGTATGTGTGTCAGGGGGCTCCAAAGCTCACTGTCAATGAGTCCAAATGGAGTGGATGA 8983
Qy 2940 GluIleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhe 2959
Db 8984 GAAGTCCCTGTCTGTAAACAGCTACTGTGTCTCTCTCCCGGACCTTCCCGAGGCTTC 9043
Qy 2960 ProAsnGlyPheSerPheHisGlyHisIleGlnTrpGlnCysPheProGlyTrpYr 2979
Db 9044 CCTAATGGCTTTCTTTTATCATGGGGCCACATACATCATCATCATCTCTTACTGTGTAT 9103
Qy 2980 LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
Db 9104 AAGCTTCATGGAACCCCATCAAGAGATGGCTTCCCAATGGTTCCTTGGAGCGGAGCTCG 9163
Qy 3000 ProSerCysLeuProCysArgCysSerThrProValIleGluTrpGlyThrValAsnGly 3019
Db 9164 CCATCTCCCTACCTTGCAGGTGTCCACACCCATCATTCACAGGAGAACCATCAAGCGCA 9223
Qy 3020 ThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeu 3039
Db 9224 ACTGATTTGGGATGTGGAAGACGGTCCAGATTGATGTCTTCAAGGGCTTCAAGCTGCTT 9283
Qy 3040 GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCys 3059
Db 9284 GACTTTCTGAAATCCTCTGTATGCCAATGGCCAAATGGTCTGAC---GTCCCATCTGTGT 9340
Qy 3060 GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer 3079
Db 9341 GAGCAGCTCAGTGGGGCTCTCCCAACCATCATCCAGCAATTTGTCTCTGAGGGCAGC 9400
Qy 3080 SerTrpLysGluAsnValIleThrTrpSerCysArgSerGlyTrpValIleGlnGlySer 3099
Db 9401 CTTTGGGAGGCAATGTGGTAACCTTACAGCTGCAGCCTGGCTACACCATGCAAGGTAGT 9460

QY 3100 SerAspLeuIleCysThrGluLysGlyValTyrSerGlnProTyrProValCysGluPro 3119
DB 9461 TCAGATCTGATTTGTACGGAAAGAGATATGAGCGCCCTTACCCACAGTGTGAACCC 9520
QY 3120 LeuSerCysGlySerProProValAlaAsnAlaValAlaThrGlyGluAlaHisThr 3139
DB 9521 CTGTCTGTGGACCCCAACCACTGAGCAATGTCAGTGGCAACAGGAGAGGCTCATACC 9580
QY 3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp 3159
DB 9581 TATGAAGCAAGCAAGTGAACCTCAGGTGTCTGCAAGGGTATGTGATGATTCGGATACAGAT 9640
QY 3160 ThrPheThrCysGlnLysAspGlyAlaGlyTyrPheProGluArgIleSerCysSerProLys 3179
DB 9641 ACATTCACTTCCGACGAGATGCGCATTTGGTCTCCTGGAAGATACCTGCACTCTTAA 9700
QY 3180 LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerVal 3199
DB 9701 AAATGCCCTGTGCCATCCCAACATGACAGCATACGTTTTCAGGAGATGACTTCCAGGTG 9760
QY 3200 AsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSer 3219
DB 9761 AACAGACAAGTTTCTGTGTCTATGTCAGAAAGGGTTTACCCACGAGAGTGAATGGTCA 9820
QY 3220 ValCysGlnLeuAspGlyThrTyrGluProPheSerAspGluSerCysSerProVal 3239
DB 9821 ACATGCCAGCCCGACGGTACATGGAGGACCACTTCTGTATGATGATGATGATGATGATGAT 9880
QY 3240 SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerIleThrPhe 3259
DB 9881 GTTTGTGGGCATCTGAAAGCCAGCGCATGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 9940
QY 3260 GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg 3279
DB 9941 GGAAGCACCATTGTTACCAAGTGTGCTCCCTGGCTACAAATTTAGAGGGGGAACAGGAA 10000
QY 3280 ValCysGlnLeuAsnArgGlnTyrSerGlyValAlaIleCysValLysGluThrArgCys 3299
DB 10001 ATCTGCAGAGAACAGACAGTGGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 10060
QY 3300 GluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyPro 3319
DB 10061 GAGACTCCAGCTGAGTTTCCCAATGGAGGCTGTCTGGCAAAACACCATCTGGACCC 10120
QY 3320 AsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCys 3339
DB 10121 AGCCTTCTGTCTCTGTCACAGAGGTACACCTCGAAGGGTCCCGGAGGACCACTGC 10180
QY 3340 ThrGluAsnGlyThrTyrSerHisProValProLeuLysProAsnProCysProVal 3359
DB 10181 ACTGCAATGGAACTGGAAATCACTGACTCCCTCTGCAAAACCAATCAATGCCCTGTTC 10240
QY 3360 ProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsn 3379
DB 10241 CCTTTGTGATCTCTGAGAACCGCTCTTCTGAAAAAGAGTTTATGTCGACCAAGAT 10300
QY 3380 ValSerIleLysCysArgGluGlyPheLeuLeuGluGlyHisGlyIleIleThrCysAsn 3399
DB 10301 GTATCTATCAAGTGCAGGGAAGGCTTCTGCTCAAAAGGCAATGGTGTATCATCGTCGAC 10360
QY 3400 ProAspGluThrTyrThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPro 3419
DB 10361 CTGACGACATGAGGACGACACCAATGCGAGATGTGAAAAAATCTCTGTGTCTTCCA 10420
QY 3420 AlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThr 3439
DB 10421 AGTCAGTAGAAGTCAATTTGCTCGAGGAGTGTATTAACAGTATGGGACATGATCAACC 10480
QY 3440 TyrSerCysTyrSerGlyTyrThrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsn 3459
DB 10481 TACTCTGTGTACAGTGGCTACATGCTAGAAAGGTTCCTCCGAGAGTGTTCCTAGAAAT 10540
QY 3460 GlyThrTyrThrSerProPheIleCysArgAlaValCysArgPheProCysGlnAsnGly 3479

DB 10541 GGAACATGGACACCATCTCTCTGTTGGAGAGTGTCTGTGTTCCCATGTCAGAAATGGA 10600
QY 3480 GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCys 3499
DB 10601 GGTGTCTGTCAAGCTCCAAATGCTGTCTCATGCCAGACGGCTGGATGGGACGCTCTCTGT 10660
QY 3500 GluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGln 3519
DB 10661 GAGAGCAATATGATCATCTCCCTGTTGATGGTGGCGCTGTGGCCCTTATCATG 10720
QY 3520 CysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerPro 3539
DB 10721 TGTGACTGCCCCACAGGCTGAGCTGGGTCCCGCTGTATCATAGCTACTTGTGAGTCCGCC 10780
QY 3540 CysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTyrThr 3559
DB 10781 TGTCTTAATGGCGGAATGCAATAGACCAACCGATGCCATTGTCTCTCAGCCTGGACA 10840
QY 3560 GlyHisAsnCysSerArgLysArgArgThrGly 3570
DB 10841 GGACATGATGTTCCAGGAAAGAGAGCCGGG 10873
RESULT 10
US-10-150-821-3
: Sequence 3, Application US/10150821
: Publication No. US20020192758A1
: GENERAL INFORMATION:
: APPLICANT: Welcher, Andrew A.
: APPLICANT: Elliott, Gary S.
: TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: 01017/37592
: CURRENT APPLICATION NUMBER: US/10/150,821
: CURRENT FILING DATE: 2002-05-16
: PRIOR APPLICATION NUMBER: US/09/911,842
: PRIOR FILING DATE: 2001-07-24
: PRIOR APPLICATION NUMBER: US 60/222,438
: PRIOR FILING DATE: 2000-08-01
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: Patent in version 3.0
: SEQ ID NO 3
: LENGTH: 11230
: TYPE: DNA
: ORGANISM: Mus musculus
US-10-150-821-3
Alignment Scores:
Pred. No.: 0 Length: 11230
Score: 16707.00 Matches: 2915
Percent Similarity: 89.72% Conservative: 289
Best Local Similarity: 81.63% Mismatches: 361
Query Match: 14 Indels: 6
DB: 14 Gaps: 5
US-09-977-053-4 (1-3571) x US-10-150-821-3 (1-11230)
QY 1 MetTyrProArgLeuAlaPheCysTyrTyrGlyLeuAlaLeuValSerGlyTyrAlaThr 20
DB 176 ATGTGTGTGGCGCTCGCCCTTTTGTGTGCTGGCTGTGGCACTGTGGTGTGGGCTGGACCAAC 235
QY 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
DB 236 TTCACGCCGTGGCCCTCTGCTCACTTCACTTCCGCTGTTCCTCGAGGCTCTCCG 295
QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
DB 296 GGGGCTCTGGGACAGCTGGCGGTACCTCCCGCTCCAGTCCAGTGGAGGAGGAGGAGGAGG 355
QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGluLeuSer 79
DB 356 AAGTGGAGCGCTGGCGCGCGGTTTCGGAGCGCGTTCGGGAGCTGCGGAGCTCAGC 415

QY 80 GluArgLeuGluLeuValPheLeuValAspSerSerValGlyGluValAsnPhe 99
DB 416 GGCACCTGGAGCTGGTCTCTCTGGTGGACGAGTCTCCAGCTGGGCGCAACCACTTC 475
QY 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValProThrAla 119
DB 476 CTCACGAGCTCAAGTTCGTGGCGCAAGCTGTCTCGAGCTTCCCGTGGTCCAGCGCC 535
QY 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
DB 536 ACGCGTGGCCATCGTCACTTCTCATCCAAAGAACAACTGGTGGCGCGCGTGAATTAC 595
QY 140 IleSerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAla 159
DB 596 ATCTCCACAGCCGCGCGCACCAACAGTGGCGCTGTCTCAGCGCGAGATCCGCGCC 655
QY 160 IleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIle 179
DB 656 ATCACTACCGGGTGGTGGCACTTATACCAAGGGGGCTTCCAGCAAGCCGCGCAATC 715
QY 180 LeuLeuHisAlaArgLysSerThrLysValValPheLeuIleThrAspGlyTyrSer 199
DB 716 CTTCGTCTCTAGAGAACTCCACCAAGTCAATATTTCTCATCCGACGCGCTATTC 775
QY 200 AsnGlyGlyAspProArgProIleAlaLeuSerLeuArgAspSerGlyValGluIlePhe 219
DB 776 AATGGCGGAGACCCAGACCTATTGCGCATCTCGCTTCGGGATTTCCGGAGTGAGATCTTC 835
QY 220 ThrPheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
DB 836 AGTTTCGGGATTTGGCAGGGGATATCCGGGAATGGAATGATGCTTCCACCCCGAG 895
QY 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
DB 896 GAAGAACAATGTTACTCTCTCCACAGTTTGAAGAATTTAGAGCTTTAGCTCCGAGGGCG 955
QY 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
DB 956 TTGCATGAAGATCTACTCTTCGGAGTTTATTCGAAGAGATGATGCCCACTGCTCTTAT 1015
QY 280 LeuCysAspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHis 299
DB 1016 CTCCTGAGGCTGGGAAGACTGCTGTGACAGATGGCCAGCTGCAAAATGTGGACACAC 1075
QY 300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyAsGlyLeuGlnTyrGlu 319
DB 1076 ACGGGTCAATTTCAATGCACTCTGTGAGAAGGGCTATTACGGGAAGGCTCTGCAGCATGAG 1135
QY 320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
DB 1136 TGCAAGCTTGGCCATCAGGGACATATAAGCCGAGACTTCTCCAGGAGATCAGCAC 1195
QY 340 CysIleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAsp 359
DB 1196 TGCAATCCCATGCTCTGACGTAAAGCACACCTCCCACTGGGAAGCACTTCCCTCGAAGAC 1255
QY 360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
DB 1256 TCGGTGTGCGAGAGGGATACCAAGATCTGGCCAGACCTGTGAGGTGTCTCCATGTCTCT 1315
QY 380 AlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
DB 1316 GCCCTGAAGCCCTCTGAAATGGTTTTTTTATACAAACACTTGCAAAACACTTCTCAAT 1375
QY 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu 419
DB 1376 GCGGCTGGGGTCCGATGTCGCGCGGGCTTGACCTTTGGGAAGCAGCATCCATTG 1435
QY 420 CysLeuProAsnGlyLeuTyrPheSerGlySerGluSerTyrCysArgValArgThrCysPro 439
DB 1436 TGTCAACCCCAATGTTGTGTGTGGACAGAAAGCTTCTGCAGAGTGAAGACGTGCCCC 1495
QY 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys 459

DB 1496 CACCTCCGACAGCCCAACACGCGCCACATCAGCTGTCTCACTCGCGGAATGTCTTACAC 1555
QY 460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr 479
DB 1556 ACCCTGT 1615
QY 480 CysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSer 499
DB 1616 TGTCAAGGAATATGCGCAGTGGATGGCCAGAGCCCCGGTGTGTAGAACCCATTGTGCC 1675
QY 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
DB 1676 ACCTTCCAGAACCCCAAGGCGTCACTTCTCCACCCAGCTGCGCAGCAGCCGCGCC 1735
QY 520 LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
DB 1736 AGGCGTGGATGACCTGTCAAGTGTCCGCGCAGGATACATTTTATCTCGGGGTGAGA 1795
QY 540 GluMetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaValCys 559
DB 1796 GAA---GTGAGATGTGCCACATCTGGGAAGTGGAGTGCCAAAGTTCAGACAGCTGTGTGC 1852
QY 560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu 579
DB 1853 AAGATGTGGAGCTCCACAAATCAGCTGTCCAAATGACATTCAGGCAAGACTGGGGAG 1912
QY 580 GlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyGlu 599
DB 1913 CAGCAGGACTCTGTAAATGCCACTGTGGCAAGTCCCAACAGCTAAAGACAACTCTGTGTGA 1972
QY 600 LysValSerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAsp 619
DB 1973 AAGGTGTAGTCCAGTCCAGCCAGCTTTACCCACCTTTACCTTCTCCAAATTTGAGAC 2032
QY 620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
DB 2033 GTGGCCATCACCTACACGCAACCGACTCATCGGTAAACCAAGCAGCTGCACCTTCTAC 2092
QY 640 IleLysValIleAspAlaGluProProValIleAspTyrCysArgSerProProVal 659
DB 2093 ATTAAGTCTAATGTGTGAACCGCTGTATAGATTGGTGGCGATCTCCACCTCCAACT 2152
QY 660 GlnValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSer 679
DB 2153 CAGGTCTGATAGAGAGAGACCCCTGCAGCTGGAGTGGAGCTCAGTTCTCAGACAACTCC 2212
QY 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
DB 2213 GGGGCTGAATTTGTCATTACACAGCAGTCCACACAGGCGACATGTTCTCATGGGAA 2272
QY 700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
DB 2273 ACGGT 2332
QY 720 IleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739
DB 2333 ATGTGATAAAGGTCTCTCCCTGTGAGGTCCCTTCAACCTGTAAACCGGGACTTTATC 2392
QY 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
DB 2393 TGTGCCAGGATAGTGTGAGTTAACTGTAGCTGAGCTGCAAGGAGGCTATGATTTTC 2452
QY 760 ThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThr 779
DB 2453 ACAGAAGGTCTCCTGAGAGTACTACTGTCTTTTGAAGATGTGTATCTCGAGACCA 2512
QY 780 TyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSer 799
DB 2513 TACTCTACAGATGGCAGACTGTGTGTATAAACGTTTTGCAAAACCATGTGTTCAAGTCC 2572
QY 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819

Db	2573	TTTGAAATGCTATACAAAACCACTCGCTGTGATGACATGATCTGTTTAAAGAGTTTCT	2632
Qy	820	GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp	839
Db	2633	GCAGCATTTGAGACTACCTCGGGAAACATGCTCCGCTCTTTGTAACGATGCTGATGAC	2692
Qy	840	IleAspCysArgLeuGluGluAsnLeuThrLysLysValCysLeuGluTyrAsnTyrAsp	859
Db	2693	ATTGACTGCGAGCTGGAGGAC---CTGACCAAAAATACTGATCGAGTATATAATACAC	2749
Qy	860	TyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyr	879
Db	2750	TATGAAAATGCTTTGCAATTTGGACACGAGGCTGGGGTGGAGCAACAGGCTGATATAT	2809
Qy	880	SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys	899
Db	2810	TCCTACCATCTTCCTGGATGTTGTACAGGAACACCCCGATGGGGCAAGGCCACGA	2869
Qy	900	SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn	919
Db	2870	TGTCACGGATTAAGAAGACTGCTCCATTTGTCGACCCCAAAATTCAGCTAATTTTAA	2929
Qy	920	IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln	939
Db	2930	ATCACAGCTAGCGTCCACTCCACAGAGAAAGAAACGATACCTTTGAATGGAGAAATC	2989
Qy	940	GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys	959
Db	2990	CAGCGACTCATTAAGACTTTGGAACCAATCAACCAATCGCTGGAAGACACCTTGAATAA	3049
Qy	960	AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeu	979
Db	3050	GAGCCCATGATTTCTTCCAGCTCGCTCGGAACAGTGTGGTGTGACAGCAATTCCTC	3109
Qy	980	GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys	999
Db	3110	GAACAGAAAGGCTTTTCTCTCTGACAGACAGGCTCTGTGCTGAGGGGGCATGTGT	3169
Qy	1000	ValAsnCysProLeuGlyThrTyrTyrAsnLeuGluIlePheThrCysGluSerCysArg	1019
Db	3170	GTCAACTGCCCCCTGGGAACCTCTTACTCTCTGGAGCATTTCCACTGTGAAGCTGCTC	3229
Qy	1020	IleGlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMet	1039
Db	3230	ATGGATCTTACCAAGATGAAGAAGGCGAGCTGGAATGCAAGCTCTGTCCCCCAAGACT	3289
Qy	1040	TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly	1059
Db	3290	CACGGGAATACCTCCATTCAGAAGCGTCTCTGAATGCAAGCTCAGTGTGAAGCAAGGC	3349
Qy	1060	ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro	1079
Db	3350	ACCTACTCTTCCAGTGGGCTGGAGACTGGGAATGCTGCTCGCTGGGTACTTATCAACCG	3409
Qy	1080	LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly	1099
Db	3410	GAATTTGGATCCCGAGCTGCTCTATGCCAGAAACCAACCAACGCGTGAAGAAAGGA	3469
Qy	1100	AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly	1119
Db	3470	GCGGTGGACATCTCTGTGTGGAGTGCCTCGCCAGTAGGAGAAATTTCTCCCGTTCTGG	3529
Qy	1120	LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe	1139
Db	3530	CTAACACCCCTGCTACCTTGGCCCTCGAGACTATTACCAACCAATGAGGGAAGTCTTTC	3589
Qy	1140	CysLeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGlu	1159
Db	3590	TGCTCTGCTTGTCTTTATGNACTTACAACCATCATCTGCGGCCACGCTCCATCACAGAC	3649
Qy	1160	CysSerSerPheSerThrPheSerAlaAlaGluSerValValProAlaSer	1179
Db	3650	TGCTCAAGTTTAGCTCTACTTTCTCAGCAGCAGAAAGACATGATGTGCCCTCGTGGCC	3709
Qy	1180	LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe	1199
Db	3710	CCTGGACATTTCCAGAACACAGTAGAGTACAGCTCAGGCTCTTTTACAGAAATGCTTTTA	3769
Qy	1200	AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCys	1219
Db	3770	AACCCCTGCCACACACAGTGGAACTGCCAACAGCTTGGGCGTGGTTATGCTCTCTGCG	3829
Qy	1220	ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro	1239
Db	3830	CCACTCTGATACACAGGCTTAAAGTGTGAACAGATATTGAATGAATGACAGCTCTCTGCT	3889
Qy	1240	CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer	1259
Db	3890	TGCTCTAATGTGGAAATTTAGAGACCAAGTTGGGGATTTCACTGCGAATGTTCTTATG	3949
Qy	1260	GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu	1279
Db	3950	GGCTATTTCAGGTCAAAATATATGTAAGAAATATATAATGAGTGTATCTCCAGCCCTTGT	4009
Qy	1280	AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe	1299
Db	4010	AATAAGGAACCTGCACTGACGCTTGGCAAGCTACCGCTGTACCTGTGTGAAGAGATAC	4069
Qy	1300	ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn	1319
Db	4070	ATGGGTGTGCTACTGTGAACACAGCTGTANAGATGCCAGTCAAGCCCTCTGCTTAACAC	4129
Qy	1320	AlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGly	1339
Db	4130	GCAGTTGTAAAGACCAAGTTTGGGGGTTCTCGTGCAAAATGCCACCCCGATTTTGGGT	4189
Qy	1340	ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr	1359
Db	4190	ACTCGGTGTGAAAAAATGTGGATGAGTGTCTCACTGACGCAATGCCAAATTCAGGCCACT	4249
Qy	1360	CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis	1379
Db	4250	TGTANGATGTGCAACAGCTTCAAGTGTCAATGTCAGCAGGCTTACAGGGACACAC	4309
Qy	1380	CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal	1399
Db	4310	TGTCAACTGAACATCAACGAGTGTGCTCCAAACCGTGTAGGAACCAAGGCCACTGTGTG	4369
Qy	1400	AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu	1419
Db	4370	GATGAACCTAAACTCATACAGTTGTAAATGTCCAGCCAGGATTTTCAGGCCACAGGTGTGAG	4429
Qy	1420	ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal	1439
Db	4430	ACAGAACAGCTTCCGGTTTAACTGGATTTTGAAGTTTCTGGCATCTACGGGTACGTC	4489
Qy	1440	MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSer	1459
Db	4490	CTGCTAGATGGAGTGTGCCAACCTCCATCCGCTAACCTGCGCATTTCTGGATGAATCC	4549
Qy	1460	SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn	1479
Db	4550	TCTGATGTCACTCACTAGGAGACCCCATCTCTATGCACTTGGAGGATGACAAAGACAC	4609
Qy	1480	ThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGlyLysIle	1499
Db	4610	ACCTTCTCTGACTGATTAACAGGCTGGTGGTCTTTATGTAATGGAAGGAAGATC	4669
Qy	1500	ThrAsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSer	1519
Db	4670	ACCAACTGCCCTCCGTAATGATGGCAATTTGGCATCATATTGCAATCACATGGAACAGT	4729
Qy	1520	AlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyAlaGlyLeu	1539
Db	4730	ATTGGTGGAGCTGGAGGCTTATATAGATGGGAATTTATCTGAGGGTGGTACTGGCCCTC	4789

Qy	1540	SerValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluAspLys	1559
Db	4790	TCCATTGGCAAAAGCCATACCTGGTGGCGGTGCATTAGCTTCTGGCGCAAGACGACAA	4849
Qy	1560	LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu	1579
Db	4850	AAAGGAGAGGGGTTCAACCCCGCTGAGTCTTTTGTGGGCTCCATAGACCGAGCTCAACCTC	4909
Qy	1580	TrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu	1599
Db	4910	TGGGACTATGTCCTGTCTCCACACGAGGTGAAGTTGTGCGCCAGCTCTCTGCCACGAGGA	4969
Qy	1600	LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyLleValGlyLysVal	1619
Db	4970	CTGAGTCGGGAAACGTGTTAGCATGGCCCGATTCTCTGTGCGGAATCACGGGGAAGGTG	5029
Qy	1620	LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro	1639
Db	5030	AAGGTGATTCCACGACGATGTTCTGCTGTGATTGTCCTTTTAGAAGGATCCGTGCTC	5089
Qy	1640	HisLeuArgThrAlaSerGluAsnLeuLysProGlySerLysValAsnLeuPheCysAsp	1659
Db	5090	CACCTGAGACCTGCATCAGGAAATCGAAGCCAGGCTCCAAAGTCAGTCTGTTCTGTGAT	5149
Qy	1660	ProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThr	1679
Db	5150	CCGGCTTCCAGATGGTTGGGAATCCTGTGCAGTATTGCTTGAACCAAGGCGAGTGCACA	5209
Qy	1680	GlnProLeuProHisCysGluArgLleSerCysGlyValProProLeuGluAsnGly	1699
Db	5210	CAACCACTCCCCCACTGTGAACGATTTGGCTGTGGGCTGCTCCCGCTTTGGAGAAATGGC	5269
Qy	1700	PheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGly	1719
Db	5270	TTCTACTCAGCGAGACTTCCATGCGGGCAGACGGTGACCTATCAGTGACCAACGATGGC	5329
Qy	1720	TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyVal	1739
Db	5330	TACTACCTGCTGGGTGATTCCGAATGTTCTGTCGACAGCAACGGGAGCTGCAACGGCATT	5389
Qy	1740	SerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisLeu	1759
Db	5390	TCACCATCTGCTCGATGTCGATGAGTGCATGCGCTCGGACTGTAGTGACACGCC	5449
Qy	1760	SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp	1779
Db	5450	TCCTGCTCCGAACACCAACCGGATCTCATGATGCTCTGTGTACCCACCATACACGGGAGAT	5509
Qy	1780	GlyLysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHis	1799
Db	5510	GGGAAAAACTGTGCAGAACCTGTAAAAATGTAAGGCTCCAGAAAAATCAGAAAAATGGCCGC	5569
Qy	1800	SerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyr	1819
Db	5570	TCCTTCTGCGCAGATTTACACCGTGGGTACTCGCATGCATTTTCTGTGACGAAGGGCAC	5629
Qy	1820	GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIle	1839
Db	5630	GAGCTGTGGGAGTGAGGACCATCACGTGTTTGGAGACTGGCGAGTGGGATCGCTCAGG	5689
Qy	1840	ProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGlu	1859
Db	5690	CGGTCTGTGAAGCCATTTCTGTGTGCTGCCACCTGTTCTCGAAAAATGGTGGTGTGAC	5749
Qy	1860	GluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeu	1879
Db	5750	GGGTGCGGATTTACATATGGCAGTAAGGTGGTGTACAGGTGTGTATGAAGGATATACTTTG	5809
Qy	1880	AlaGlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProVal	1899
Db	5810	TCTGGGGATGAAGACTGACATGCTTCTAGTGGTTCCTGGAGTCAATTCCTCTCTCTGTG	5869
Qy	1900	CysGluProValLysCysSerSerProGluAsnLeuAsnAsnGlyLysTyrIleLeuSer	1919

[illegible]

6950 ACCCTCTCAACTGTGGGAAACCCCTCCCAATTACAGATGGCTTTTTCGAAAGGAGAAAGC 7009
2280 PheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAsp 2299
7010 TTTGAAGTAGGCTCAAGGTTTCAGTTTGTCTGTAATGAGGGATATGAGCTCGTGTGGTAT 7069
2300 SerSerTrpThrCysGlnLysSerGlyLysValTrpAsnLysLysSerAsnProLysCysMet 2319
7070 AATCTTGACCTGCGCAAAATCTGCAAAATGAGGTAAAGAGCAAGCCGGAAGTGTGTC 7129
2320 ProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThr 2339
7130 CCCACCAAGTGTGCAGAGCTCTCTCTTAGAAAAACAGCTGCTATTGAAGGAATTAGCT 7189
2340 ThrGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSer 2359
7190 TCCGAGTAGGAGTATGACCAATTCTCTGTAAAGAGGGGCGATCCCTTGCAGAGCCCTCT 7249
2360 ValLeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleVal 2379
7250 GTCTGAGTGTCTGCATCCCGGCAATGGAATGGTTCCTTCTATTGTAGATGGTC 7309
2380 LeuCysThrProProLeuLysSerPheGlyValProLeuProSerSerAlaLeuHis 2399
7310 CTTTGTCCCTCGCTCCCTTGATTCTCTCGGGTCCCTCGCTCTTCGGTCTCTTCAT 7369
2400 PheGlySerThrValLysTrpSerCysValGlyPhePheLeuArgGlyAsnSerThr 2419
7370 TTTGGCAGTACTGTCAAGTATCTGTGTGTCGACGGGTCTTCTTAAAGAGCGATGCCAAC 7429
2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
7430 ATCTCTGCGCAGCTGATAGCACTGGAGTCTTCATTTGCCGGAATGGCTTCGGTAGAA 7489
2440 CysProGlnProGluGluLeuProAsnGlyIleLeuAspValGlnGlyLeuAlaLysLeu 2459
7490 TGTCCCAACCTGAGGAGATCTCAACGGTATCATCCAGCTACAGGGCTTGCTATCTC 7549
2460 SerThrAlaLeuThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeu 2479
7550 AGCACCAACCTCTACACCTGCAAGCAGGCTTTGAGTTAGTGGGCAATGCTACACCCCTC 7609
2480 CysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeu 2499
7610 TGTGGGAAAAATGGCCAGTGGCTCGAGGGAACCAATGTGCACAAACCAATTGAATGCCCA 7669
2500 LysProLysGluLeuLeuAsnGlyLysPheSerTrpThrAspLeuHisTrpGlyGlnThr 2519
7670 GAGCCCAAGGAGATTTTAAATGGCCAAATCTCTTCCTCGTGGAGCTTTCAGTATGACAAAC 7729
2520 ValThrTrpSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
7730 ATCACATATCTTTGTGACCGGGCTTCGGCTCGAAGTCCGAAGTCCCAATCCCTGACCTGTTA 7789
2540 GluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerPro 2559
7790 GAGACAGGTGACTGGGATATGATCCCTCTCTGTGATGCCATCCACTGCAGTGACCCA 7849
2560 GlnProIleGluAsnGlyPheValGluGlyAlaAspTrpSerTrpGlyAlaIleIle 2579
7850 CAGCCCAATTGAAATGGTTTCGTAGAAGGTGCGGATTTACAGATACGCTGCGATGATCATC 7909
2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
7910 TATAGTGTCTTCTCCCTGGGTTTCAGGTGCTTGGTTCATGCCATGACAGCTGTGAAGATCG 7969
2600 GlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIle 2619
7970 GGAATGTCAAGCTCCAGCCCAACTGTGTACCCATAGACTGCGGTCTCCCTCTCCATATA 8029
2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTrpPheGluGlnLysAsp 2639
8030 GACTTTGGTGTACTTAAGTCAGAGATGGCCAGGACATTTTGTATCAAGAGATGAC 8089

2640 MetMetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLys 2659
8090 ATGATGGAAGTCCATATCTGGCT-----CACCTCAACATTTGGAAGCAACAGCTAAG 8143
2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThr 2679
8144 GCCTTGGAAAAATACAAAGGAGTGCCTGCTGCATGCATGCCATCCCTCTCTATGGCAGC 8203
2680 MetValSerTrpThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCys 2699
8204 ATGTTTCTTACAGCTGCGAGCTGTGTATGAATGCTTGGGAATCCCTGTGCTGATCTGC 8263
2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
8264 CAGAAAGATGTTAGTGAATGTTGCCGACCTCTTTCATTTCCATTAATGATGATTTG 8323
2720 ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
8324 CCTGTTGCTCCCGAAAAATGGCTTTTACATTTTACACAGACGACTATGGCGAGTGTGCA 8383
2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
8384 CAATATAGCTGCAAGCGCGGCACATTTCTAGAAGGCTCCCACTTAAGACTCTGTCTGCAG 8443
2760 AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysAlaLysLysProAsn 2779
8444 AATAGCAGTGGAGTGGCACTGTTCCAGCTGTGAAGCCATCTCAGTGAAGTGAAGCAAC 8503
2780 ProValMetAsnGlySerIleLysGlySerAsnTrpThrTyrLeuSerThrLeuTyrTrp 2799
8504 CCATCTGGGAATGATCCATCAAGGAGATGACTACTCTTACCTGCGGTGTGTATCTATAC 8563
2800 GluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspLys 2819
8564 GAGTGTGACTCTGCTATATTCTCAATGGCTTAAGAAGAGGACATGCCAAGAAATAGA 8623
2820 AsnTrpAspGluAspGluProIleCysValleProValAspCysSerSerProProValSer 2839
8624 GATTGGATGGCATGAGCCCATGTGTATCTGTAGACTGTGCTCACCCCAAGTCCCC 8683
2840 AlaAsnGlyGlnValArgLysAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
8684 ACCAATGGCGAGTGAAGGAGGAAGAATACACATTTCCAAAGGAGATATACATCTCTTGC 8743
2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
8744 CGTGAAGGGTTCTACTTGAAGGAGCGCAGGAGTGTATCTGTCTTACCAATGGAAGTTG 8803
2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsn 2899
8804 AGTGGTGCCACTCCAGCTGCATGCTGTGTAGATGCTCTGCCCCACACAGGTGCCAAT 8863
2900 GlyValThrGluGlyLeuAspTrpGlyPheMetLysGluValThrPheHisCysHisGlu 2919
8864 GGGGTGCGAGTGGCTTAGACTATGGCTTCAAGAAGGAGTAGGCTTCCACTGTCTAGAG 8923
2920 GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
8924 GGTATGTGTCGAGGGGCTTCAAGACTCACCTGTGCTGCTCAATGGAGCTTGGGATGCA 8983
2940 GluIleProLeuLysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhe 2959
8984 GAAGTCCCTGTCTGTAACCCAGCTACTCTGTGTCTCTCTGCCGACTTCCCCAGGGCTTC 9043
2960 ProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTrp 2979
9044 CCTAATGGCTTTCTTTTATCATGSGGGCCACATACAGTATCAGTGTCTTACTGTGTTAT 9103
2980 LysLeuHisGlyAsnSerSerArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
9104 AAGCTTCATGGAAACCCATCAAGAAGATGCTTCCCAATGGTCTCTGGAGCGGAGCTCG 9163

; PRIOR APPLICATION NUMBER: 60/218,591
 ; PRIOR FILING DATE: 2000-07-17
 ; PRIOR APPLICATION NUMBER: 60/271,814
 ; PRIOR FILING DATE: 2001-02-27
 ; PRIOR APPLICATION NUMBER: 60/215,855
 ; PRIOR FILING DATE: 2000-07-03
 ; PRIOR APPLICATION NUMBER: 09/839,446
 ; PRIOR FILING DATE: 2001-04-19
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 21

; LENGTH: 6153
 ; TYPE: DNA
 ; ORGANISM: Unknown Organism
 ; FEATURE:
 ; NAME/KEY: modified_base
 ; LOCATION: (1010)
 ; OTHER INFORMATION: a, c, t, g, other or unknown
 ; DESCRIPTION: Description of Unknown Organism: POLYX
 ; OTHER INFORMATION: 10327789_1
 ; US-09-898-570-21

Alignment Scores:
 Pred. No.: 0 Length: 6153
 Score: 8427.00 Matches: 1535
 Percent Similarity: 78.65% Conservative: 1
 Best Local Similarity: 78.60% Mismatches: 5
 Query Match: 42.13% Indels: 412
 DB: 9 Gaps: 3

US-09-977-053-4 (1-3571) x US-09-898-570-21 (1-6153)

QY 1993 TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSer 2012
 DB 193 TATATCTTCGTGGCTTTCAGACCAATGTAATGCCCTGGCCGACGCGCAAGTGGAGTAGAAGT 252
 QY 2013 AspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerPro 2032
 DB 253 GACCAGCAGTGGCTGCTCTCTCTGATGATGACCCACCCATTTGTGGACCCAGCCCTCTCCA 312
 QY 2033 GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSer 2052
 DB 313 GAGACTGCCCATCGGCTCTTTGGAGACATTTGCAATCTACTACTGCTCTGATGGTTACAGC 372
 QY 2053 LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGly 2072
 DB 373 CTAGCAGCAATTCACGCTCTCTGCAATGCCAGGCGAGTGGTATGACCCCAAGGT 432
 QY 2073 GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer 2092
 DB 433 CAAGACATGCCCGTTGTATAGCTCATTCTGTGAAACACCTCCATCGGTTTCTCTATAGC 492
 QY 2093 IleLeuGluSerValSerLysAlaLysPheAlaGlySerValValSerPheLysCys 2112
 DB 493 ATCTTGGAAATCTGTGAGCAAGCAAAATTTGCAGCTGGCTCAGTTGTGTGAGCTTTAAATGC 552
 QY 2113 MetGluGlyPheValIleAsnThrSerAlaLysIleGluCysMetArgGlyGlnTrp 2132
 DB 553 ATGGAGGCTTTGTACTGAACACCTCAGCAAGATTTGAATGATGAGAGTGGGAGTGG 612
 QY 2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle 2152
 DB 613 AACCTTCCCATGTCATCCAGTGCATCCCTGTGGGTGGAGAGCCACCAAGCATC 672
 QY 2153 MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
 DB 673 ATGAATGGCTATGCAAGTGGATCAACATACAGTTTTCGAGCCATGCTGGCTTACAGCTGC 732
 QY 2173 AsnLysGlyPheTyrIleLysGlyGlnLysLysSerThrCysGluAlaThrGlyGlnTrp 2192
 DB 733 AACAGGGGGTTCTACATCAAGGGGAAAGAGAGACCTCGGAAGCCACAGGGGAGTGG 792
 QY 2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212

DB 793 AGTAGTCTTATACCCGACGTCGCCACCCGGTATCTTGGTGAACCCACTAAGTTGAGAT 852
 QY 2213 GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn 2232
 DB 853 GGCCTTTCTGGAGCATACAACTGGCAGGATCTTTTGAGAGTGAAGTGAAGTATCAGTGTAAAC 912
 QY 2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHis 2252
 DB 913 CCGGGCTATAAGTCAGTCGGAGTCTCTGTATTTGTCTGCCAAGCCAATCGCCACTGGCAC 972
 QY 2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIleGlnAsn 2272
 DB 973 AGTGAATCCCTCTGATGTGTCTCTCTGCACTGTGGAACCTCTCCCGATCCAGAT 1032
 QY 2273 GlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
 DB 1033 GGCCTTCATGAAGAGGAGAAACCTTTGAAGTAGGGTCCAAGGGTCAGTTTTTCTGTAAATGAA 1092
 QY 2293 Gly---TyrGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsn 2311
 DB 1093 GGGTTATNGAGCTTTGTTGGGACAGTTCTTGGACATGTCAGAAATCTGCAAAATGGAAT 1152
 QY 2312 LysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsn 2331
 DB 1153 AAGAAGTCAAAATCCAAAGTGCATGCTGCCAAGTCCAGAGCGGCCCTCTTTGGAAAC 1212
 QY 2332 GlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGlu 2351
 DB 1213 CAGCTAGTATTAAAGGAGTTGACCCAGGAGTAGGAGTTGTGACATTTTCTCTGTAAGAA 1272
 QY 2352 GlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAsp 2371
 DB 1273 AGGCATGCTCTGCAGGCCCTCTGCTCTGTAATGCTTGCCATCCAGCAATGGAATGAC 1332
 QY 2372 SerPheProValCysLysIleValLeuCysThrProProLeuIleSerPheGlyVal 2391
 DB 1333 TCTTTCCTCTGTTTGAAGATGTTCTTGTGTACCCACCTCCCTTAATTTCTTGTGTGTGTC 1392
 QY 2392 ProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGly 2411
 DB 1393 CCATTCCTCTCTCTGCTCTTCAATTTTGAAGTACTGTCAAGTATTTCTGTGTAGTGGG 1452
 QY 2412 PhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerPro 2431
 DB 1453 TTTTCTCTAAGAGAAATTTCTACACCTCTGCCAACCCTGATGGACCTGGAGCTCTCCA 1512
 QY 2432 LeuProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIleIle 2451
 DB 1513 CTGCCAGAAATGTTCTCAGTAGAATGTCCCAACCTGAGGAAATCCCAATGGAATCAT 1572
 QY 2452 AspValGlnGlyLeuAlaTyrIleuSerThrAlaLeuTyrThrCysLysProGlyPheGlu 2471
 DB 1573 GATGTCAAGGCTTGGCTTATCTCAGCAGCAGCTCTCTATACCTGCAAGCAGGCTTTGAA 1632
 QY 2472 LeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyLysPro 2491
 DB 1633 TTGTGGGAAATACTACACCTTTGTGGAGAAATGGTCACTGGCTTGGAGGAAACCA 1692
 QY 2492 ThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyr 2511
 DB 1693 ACATGTAAAGCAATTCAGTGCCTGAAACCCCAAGAGATTTTGAATGGCAATTTCTCTTAC 1752
 QY 2512 ThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGlu 2531
 DB 1753 ACGACCTTACATATGACAGACCGTTTACTCTTGTCAACCGAGGCTTTTGGCTCGAA 1812
 QY 2532 GlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCys 2551
 DB 1813 GGTCCAGTGCCTTGGACCTGTTTAGACAGAGGTGATGGATGTAGTGTGATGCTTGTTC 1872
 QY 2552 AsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAsp 2571

3274 ----- 3274
Db CTGAGTCAGGTCGGAGTTGGCTAGAGGTCCATCCAGACCCCTGTTTGCTGGGTATCA 4152
3274 ----- 3274
Db GCAGCAGAGGTGCAGAAACAGCGATATGTGTGAACCGCAATGCTGCTGCTGATCAAT 4212
3274 ----- 3274
Db CCTCTGAAGTTTTGTCTCAGAGAAATACCCGGCCATGTGAGGTGTGCTGCGCCCTAC 4272
3275 ----- GlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAla 3292
Db TGGGGGGGAACAGGGAACGTGTCTCCAGGAGAACACAGAGTGGAGTGGGGTGGCA 4332
3293 IleCysIysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIle 3312
Db ATATGCAAGAGACACAGTGTGTAACCTCCACTTGAATTTCTCAATGGGAAGCTGACAT 4392
3313 GluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGlu 3332
Db GAAACAGGACGACTGACCCCAACGTGTATATTCCTGCAACAGAGGTACAGTCTTGA 4452
3333 GlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCys 3352
Db GGGCCATCTGAGGCACACTGCACAGAAATGGAACCTGGAGCCACCCAGTCCCTCTGCG 4512
3353 LysProAsnProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLys 3372
Db AAACCAATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4572
3373 GluPheTyrValAspGlnAsnValSerIleIysCysArgGluGlyPheLeuLeuGlnGly 3392
Db GAGTTTATGTGTATCAAGATGTGCTCATCAATGTAGGAGAGTTTCTGCTGCGAGGC 4632
3393 HisGlyIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGlu 3412
Db CACGGCATATTACCTGCACCCCGACGAGAGCTGGACACAGACAAAGGCCAATGTGAA 4692
3413 Lys----- 3413
Db AGAAGATATACACACAGCCCAAGTCCCTGAAATTTTCAGCTAGCAGCTTATTGCAAT 4752
3413 ----- 3413
Db AGAATGTTTATTTTGGGGGAGGGGTTCAAGATGGCCCACTAGAAACAGCTGTGGCCGGA 4812
3413 ----- 3413
Db GCCTCCACCGAGAGAACAAACAAACAAAGCGAGAAAAAGCAGGTGTGTACAAACGGCCCA 4872
3413 ----- 3413
Db CTGGGAGCCATGTTGGGGCAAGCAGAGCTCCACCCCGACCAAGAGAGGTGGACCTCCC 4932
3413 ----- 3413
Db TGGGGGAATTTACAGAACTCCAGCCAGGGGTTTATGAAACAGACTCTGTATCTCCCTGAGA 4992
3413 ----- 3413
Db TGGAGCCCTGGGGCTCCATGTGGCCATGCTCTCCACAGATCAGAGGCTTAGTCTCTCC 5052
3413 ----- 3413
Db CCTGTGGCTCTGAGGAATCCAGCAGGCTGGACTAGTGGGTTCCCAACAGACAGTGT 5112
3413 ----- 3413
Db ACCTGTCTCCCAAGGGGACGTAGAGCGGCTTTGTTAAGCGAGTCCCTGATCCATGCT 5172
3413 ----- 3413

Db CCTGTTGGATGAGACCCCCACAAACAGGGGTCAACGATGAGACCCCCCAACACAGG 5232
3413 ----- 3413
Db GGTCAACAGACACCTTATACAAAGGGTGTCTGTAGCATCAGGTCACTGCCCCCTCTGG 5292
3413 ----- 3413
Db ACAGAGCTCCAGAGGAAGAGCAGGAGCCCATCTTTGCTGTTCTGACAGCTCCGCTGA 5352
3413 ----- 3413
Db AAGCACAGAAATGGGCAGAGGCTAGGATTCATGATGAATGAAGAAGTAGGTCTCAGAAG 5412
3413 ----- 3413
Db TGGGTAATAATGAAGTTGCTGAGCTAAAGGAACATGTTCTAAACCAATGCAAGACGCC 5472
3413 ----- 3413
Db AAGAACAGGATATAACATTACAGGATCCGTTAACCAAGATAACCAAGTTTAGAAGGAAT 5532
3413 ----- 3413
Db GTAAATGACCTGATGAGAGTGAATAACACACACAGAGAACTTCACATGCAACAAACAAA 5592
3413 ----- 3413
Db CAAGGCCAAACATTCAGCTTCAGGAAATCCAGAGAACCCAGTAAGTACTCCTAGAGAG 5652
3414 ----- IleSerCysGlyProProAla 3420
Db ATCAACCCCAAGACACATATCTCAGGTTCTCCCAAGAAATCTCATGTGTCACACAGCT 5712
3421 HisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
Db CACGTAGAAATGCAATTTGCTCGAGGCGGTACATTAATCAATATGGAGACATGATCACCTAC 5772
3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly 3460
Db TCATGTTACAGTGTGATACATGTTGGAGGTTTCTCGAGAGTGTGTTGTTTGAATAATGA 5832
3461 ThrTrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGly 3480
Db ACATGACATCACCTCTCTATTTCAGAGCTGTCTGTGATTTCCATGTCAGAAATGGGGC 5892
3481 IleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGlu 3500
Db ATCTGCCAACCGCCCAATGCTTGTCTCTGTCAGAGGCTGGATGGGGCGCTCTCTGTGAA 5952
3501 GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys 3520
Db GAACCAATGCAATCTCTCTGTCGAGAGGTTCTGTGTGAGGAGTCTGTGTGAGGCTCTTAC 6012
3521 AspCysProGlyTyrThrGlySerArgCysHisThr 3533
Db GACTGCCCGCTGCTGCTGCAAGGGGCTCTGCTGTCTATACA 6051

RESULT 12

US-09-839-446-21
; Sequence 21, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; CURRENT FILING DATE: 2001-04-19

;; PRIOR APPLICATION NUMBER: 60/198,293
;; PRIOR FILING DATE: 2000-04-19
;; PRIOR APPLICATION NUMBER: 60/198,645
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: 60/210,809
;; PRIOR FILING DATE: 2000-06-09
;; PRIOR APPLICATION NUMBER: 60/199,476
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: 60/200,025
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: 60/224,610
;; PRIOR FILING DATE: 2000-08-11
;; PRIOR APPLICATION NUMBER: 60/200,024
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: 60/199,880
;; PRIOR FILING DATE: 2000-04-26
;; PRIOR APPLICATION NUMBER: 60/218,591
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 60/271,814
;; PRIOR FILING DATE: 2001-02-27
;; NUMBER OF SEQ ID NOS: 49
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 21
;; LENGTH: 6153
;; TYPE: DNA
;; ORGANISM: Unknown Organism
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (1010)
;; OTHER INFORMATION: a, c, t, g, other or unknown
;; OTHER INFORMATION: Description of Unknown Organism: POLYX
;; OTHER INFORMATION: 1032789_1
US-09-839-446-21

Alignment Scores:
Pred. No.: 0 Length: 6153
Score: 8427.00 Matches: 1535
Percent Similarity: 78.65% Conservative: 1
Best Local Similarity: 78.60% Mismatches: 5
Query Match: 42.19% Indels: 412
DB: 10 Gaps: 3

US-09-977-053-4 (1-3571) x US-09-839-446-21 (1-6153)

Qy	1993	TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSer	2012
Db	193	TATACTCTTGGTCTTGACACCATTAATGCTGGCCGCGCAAGTGGAGTAGAAGT	252
Qy	2013	AspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerPro	2032
Db	253	GACCAAGCATGCTGGCTGCTCTCTGTGATGAGCCACCCATTTGTGGACCAACGCTCTCCA	312
Qy	2033	GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyLysSer	2052
Db	313	GAGACTGCCATCGGCTCTTTGGAGACATTGCAATCTACTACTGCTCTGATGGTTACAGC	372
Qy	2053	LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProGluGly	2072
Db	373	CTAGCAGACAAATTCACAGCTTCTCTGCAATGCCAGGCAAGTGGGTACCCCGAGAAGT	432
Qy	2073	GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer	2092
Db	433	CAAGCATGCCCGTGTATAGCTCATTTCTGTGAATAAACCTCCATCGGTTTCCTATAGC	492
Qy	2093	IleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys	2112
Db	493	ATCTTGGATCTGTGAGCAAGCAAAATTTGCAGCTGGCTCAGTTGTGAGCTTTAAATGC	552
Qy	2113	MetGluGlyPheValLeuLeuThrSerAlaLysIleGluCysMetArgGlyGlnTrp	2132
Db	553	ATGGAGGCTTTGTACTGAACACCTCAGCAAGAAATGTAATGATGAGGTTGGGCAGTGG	612
Qy	2133	AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle	2152

Db	613	AACCCCTTCCCATGTCATCCAGTCATCCCTGTGGGTGTGGAGACCACCAAGCATC	672
Qy	2153	MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys	2172
Db	673	ATGAATGGCTATGCAAGTGGATCAAACTACAGTTTTTGGAGCCATGGTGGCTTACAGTGC	732
Qy	2173	AsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThrGlyGlnTrp	2192
Db	733	AACAGGGGTTTTCATCATCAAGGGGAAGAGAGACCTCTGCGAGCCACAGGCGCATGG	792
Qy	2193	SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn	2212
Db	793	AGTAGTCTATACCGAGTGCACCGGTATCTTGTGTGAACACCACTAAGGTTGCAAT	852
Qy	2213	GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn	2232
Db	853	GGCTTTCTGGAGCATCAAACTGGCAGATCTTTGAGAGTGAAGTATCATGTATAC	912
Qy	2233	ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHis	2252
Db	913	CCGGGCTATAGTCAGTCGGAAGTCTGTATTTGTCTGCCAGGCCAATCCGCACTGGCAC	972
Qy	2253	SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIleGlnAsn	2272
Db	973	AGTGAATCCCTCTCATGTGTGTCTCTCTCGACTGTGGAACCTCCCGCATCCAGAT	1032
Qy	2273	GlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGlu	2292
Db	1033	GGCTTTCATGAAAGGAGAAAACCTTTGAAGTAGGGTCCAAAGGGTCAGTTTTTCTGTAAAGAA	1092
Qy	2293	Gly---TyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTrpAsn	2311
Db	1093	GGGTTATNGAGCTTTGTGGGAGCAGTTCTTGGACATGTTCAGAAATCTGGCAATGCAAT	1152
Qy	2312	LysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsn	2331
Db	1153	AAGAGTCAAAATCCAAAGTGATGCTGCTGCCAAGTGCAGAGCGGCCCTCTTGGAAAAC	1212
Qy	2332	GlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGlu	2351
Db	1213	CAGCTAGTATTAAAGGAGTTGACCCAGGAGTAGGAGTTGTGACATTTTCTGTAAAGAA	1272
Qy	2352	GlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAsp	2371
Db	1273	AGGCATGCTCTGCAAGCCCTCTGTCTGGAATGTCTGCCATCCAGCAATGGAATGAC	1332
Qy	2372	SerPheProValCysLysIleValLeuCysThrProProProLeuLysSerPheGlyVal	2391
Db	1333	TCTTTCCCTGTTTGAAGATTGTTCTTTGTACCCACCTCCCTTAATTTCTCTTGGTGTCTC	1392
Qy	2392	ProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGly	2411
Db	1393	CCCATCTCTTCTGCTCTTCTATTTTGAAGTACTGTCAAGTATTTCTGTGTAGGTGGG	1452
Qy	2412	PhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerPro	2431
Db	1453	TTTTTCTTAAGAGGAATTTCTACCACTCTGCAACCTGTGCACTGTGCACTGAGCTCTCCA	1512
Qy	2432	LeuProGluCysValProValGluCysProGlnProGluGluLeuProAsnGlyIleIle	2451
Db	1513	CTGCCAGATGTTGTTCCAGTAGAATGTCCCAACCTGAGGAAATCCCAATGGAATCATT	1572
Qy	2452	AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGlu	2471
Db	1573	GATGTGCAAGGCTTCCCTATCTCAGCAGACGCTCTCTATACCTGCAAGCCAGGCTTTGAA	1632
Qy	2472	LeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysPro	2491
Db	1633	TTGGTGGGAAATCTACCACTTTTGTGGAGAAATGGTCACTGGCTTGGAGGAAACCA	1692
Qy	2492	ThrCysLysAlaIleGluCysLeuLysProLysGluLeuLeuAsnGlyLysPheSerTyr	2511

1693 ACATGTAAAGCCATTGAGTGGCTGAAACCCAGGAGATTTTGAATGCAAAATCTCTTAC 1752
2512 ThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGlu 2531
1753 ACGGACCTACACTATGACACAGACCGTTACCTACTCTTGCACACGAGGCTTTCCGGCTCGAA 1812
2532 GlyProSerAlaLeuThrCysLeuGluThrGlyAspTTPAspValAspAlaProSerCys 2551
1813 GGTCCACAGTGGCTTGGACCTGTTTAGAGACAGGTTGATTTGGATGTAGATGCCCATCTTGC 1872
2552 AsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAsp 2571
1873 AATGCCATCCACCTGTGATTTCCCCACACCCCATTTGAAAATGTTTGTAGAGGTGCAGAT 1932
2572 TyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHis 2591
1933 TACAGCTATGGTCCATAATCATCTACAGTTGCTTCCCTGGGTTTCAGGTGGCTGGTCAAT 1992
2592 AlaMetGlnThrCysGluGluSerGlyTTPSerSerSerIleProThrCysMetProIle 2611
1993 GCCATGCAGACCTGTGAAGAGTCAGATGGTCAAGTTCCATCCCAACATGATATGCCAATA 2052
2612 AspCysGlyLeuProProHisIleAspPheGlyAspCysThrIysLeuIysAspAspGln 2631
2053 GACTGTGGCTCCTCTCTCATATAGATTTTGGAGACTGTACTAACTCAAAAGATGACCAG 2112
2632 GlyTyrPheGluGlnGluAspMetMetGluValIleProTyrValThrProHisProPro 2651
2113 GGATATTTTGGACACAGACGACATGATGGAAGTTCCATATGTGACTCTCCACCTCCT 2172
2652 TyrHisLeuGlyAlaValAlaIleThrTTPGluAsnThrIysGluSerProAlaThrHis 2671
2173 TATCATTTGGAGAGTGGCTGAAACCTGGGAAAATACAAAGGAGTCTCTGCTACACAT 2232
2672 SerSerAspPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeu 2691
2233 TCATCAAACTTCTGTATGGTACCATGGTTTTCATACACCTGTAAATCCAGGATGAACTT 2292
2692 LeuGlyAspProValLeuIleCysGlnGluAspGlyThrTTPAsnGlySerAlaProSer 2711
2293 CTGGGGAACCTCTGCTGATCTGCCAGGAGATGGAATTTGGAAATGGCAGTGACCATCC 2352
2712 CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
2353 TGCATTTCAATGAAATGAGCTTGCCTACTGCTCTCTGAAAATGGCTTTTGGCGTTTACA 2412
2732 GluThrSerMetGlySerAlaValGlnTyrSerCysIysProGlyHisIleLeuAlaGly 2751
2413 GAGACTAGCATGGGAAGTCTGTGCAGATAGCTGTAAACCTGGACACACATTTCTAGCAGGC 2472
2752 SerAspLeuArgLeuCysLeuGluAsnArgIleTTPSerGlyAlaSerProArgCysGlu 2771
2473 TCTGACTTAAGGCTTTGTCTAGAGAAATAGAAAATGGAGTGGTGGCTCCCCACGCTGTGAA 2532
2772 AlaIleSerCysIleIysProAsnProValMetAsnGlySerIleIysGlySerAsnTyr 2791
2533 GCCATTTTCATGCAAAAGCCCAATTCAGTCAATGATGATTCATCAAGAGAACCACTAC 2592
2792 ThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGlu 2811
2593 ACATACCTGACGAGTGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
2812 ArgArgThrCysGlnAspAspIysAsnTTPAspGluAspGluProIleCysIleProVal 2831
2653 AGGAGAACATGCCAGGATGACAAAATCTGGGATGAGGATGAGGCCAATTTGCAATTCCTGTG 2712
2832 AspCysSerSerProProValSerAlaAspGlyGlnValArgGlyAspGluTyrThrPhe 2851
2713 GACTGCAGTTTCAACCCCGAGTCTCAGGCCAATGGCCAGTGGAGAGACGAGTACATTC 2772
2852 GlnIysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArg 2871
2773 CAAAAGAGATTGAATACACTTGCATGTAAGGGTTCTTGTGAGGAGGAGCCAGGAGTCCG 2832

2872 ValCysLeuAlaAsnGlySerTTPSerGlyAlaThrProAspCysValProValArgCys 2891
2833 GTTGTCTTCCCAATGGAAGTTGGAGTGGAGCCACTCCCGACTGTGCTGTCCCTGTAGATGT 2892
2892 AlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetIys 2911
2893 GCCACCCCGCCACACACTGGCCCAATGGGTCAGCGAAGSCCTGGACTATGGCTTCATGAAG 2952
2912 GluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProIysLeuThrCys 2931
2953 GAAGTAACATTCCTGCTCAGCGGGCTACATCTTGCACGGTGTCTCCAAAACCTCACCTGT 3012
2932 GluSerAspGlyAsnTTPAspAlaGluIleProLeuCysIysProValAsnCysGlyPro 2951
3013 CAGTCAGATGGCAACTGGGATGAGAGATTCCTCTCTGTAAACCGATCACTGTGGACCT 3072
2952 ProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyHisIle 2971
3073 CCTGAAGATCTTCCCAATGGTTTCCCTAATGGTTTTCCTTTATTCATGGGGCCATATA 3132
2972 GlnTyrGlnCysPheProGlyTyrIysLeuHisGlyAsnSerSerArgArgCysLeuSer 2991
3133 CAGTATCAGTGTCTTCTGGTTATAGCTCCATGGAATTCATCAAGAAGGTGCTCTCC 3192
2992 AsnGlySerTTPSerGlySerSerProSerCysLeuProCysArgCysSerThrProVal 3011
3193 AATGGCTCTCGAGTGGCAGCTCAGCTTCTGCTGCTTGCAGATGTTCCACACCAGTA 3252
3012 IleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyIysAlaAlaArgIleGln 3031
3253 ATTGAATATGGAATGTCAATGGGACAGATTTTGACTGTGGAAAGGCAGCCCGGATTCCAG 3312
3032 CysPheIysGlyPheIysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln 3051
3313 TGCTTCAAGGCTTCAAGCTCTAGGACTTCTGAAATCACTGTGAAGCCGATGGCCAG 3372
3052 TTPSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIlePro 3071
3373 TGGAGCTCTGGGTTCCCCCACTGTGAACACACTTCTTGTGGTTCTCTTCCAAATGATACCA 3432
3072 AsnAlaPheIleSerGluThrSerSerTTPLeuGluAsnValIleThrTyrSerCysArg 3091
3433 AATGGCTTCATCAGTGGACCCAGCTCTTGGAGGAAAATGTGATAACTTACAGCTGCAGG 3492
3092 SerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluIysGlyValTTPSer 3111
3493 TCTGGATATGTCATACAAAGCAGTTCAGATCTGATTTGTACAGAGAAGGGGTATGGAGC 3552
3112 GlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAlaAsnAla 3131
3553 CAGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCCCCACCGTCTGTGCCCAATGCA 3612
3132 ValAlaThrGlyGluAlaHisThrTyrGluSerGluValIysLeuArgCysLeuGluGly 3151
3613 GTGGCAACTGGAGAGGCACACACTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3672
3152 TyrThrMetAspThrAspThrAspThrPheThrCysGlnIysAspGlyArgTTPPhePro 3171
3673 TATACGATGGATACAGATACAGATACATTCACCTGTGCAGAAAGATGGCTGTGGTTCCT 3732
3172 GluArgIleSerCysSerProIysCysProLeuProGluAsnIleThrHisIleLeu 3191
3733 GAGAGAAATCTCTGAGTCTCTAAAATAATGCTCTCCCGGAAAACATACATATACTT 3792
3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
3793 GTACATGGGACGATTTTCAAGTGTGAATAGCAAGTTTCTGTGTCTCATGTGCAGAAAGGTAT 3852
3212 ThrPheGluGlyValAlaHisIleSerValCysGlnLeuAspGlyThrTTPGluProPhe 3231
3853 ACCTTTGGGGAGTTAAACATATCATATGATGATGATGATGATGATGATGATGATGATGATGAT 3912


```

; Sequence 24, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Daniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Valda
; APPLICANT: LU, Dyrung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAPALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 3991
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 3998749CB1
US-10-311-623-24

Alignment Scores:
Pred. No.: 0 Length: 3991
Score: 7471.00 Matches: 1327
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 3
Query Match: 37.41% Indels: 2
DB: 17 Gaps: 0

US-09-977-053-4 (1-3571) x US-10-311-623-24 (1-3991)
QY 2205 GlyGluProProValGluAsnGlyPheLeuGluHisThrThrGlyArgPheGlu 2224
Db 1 GGTGAACCACTAAGGTGTGAAGATGGCTTTCTGGAGCATCAACTGGCAGATCTTTTGAG 60
QY 2225 SerGluValArgTyrGlnCysAsnProGlyTyrIleSerValGlySerProValPheVal 2244
Db 61 AGTGAAGTGAAGTATCAGTGTAAACCCGGCTATAAGTCAGTGGAGTCTCTGATTGTC 120
QY 2245 CysGln-AlaAsnArgHisTrpHisSerGluSerProLeuMetCysValProLeuAspCys 2264
Db 121 TGCCAAAGGCCAATGCGCCACTGGCACAGTGAATCCCTCTGATGTGTCTCTCCGACTG 180
QY 2264 sGlyLysProProGlnGlnGlnGlyPheMetLysGlyGluAsnGlyValGlyLys 2284
Db 181 TGGAAACCTCCCGCATTCAGATGGCTTATGAGTTGTTGGTGGACAGTCTTGGAGTGGTC 240
QY 2284 rLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTrpThrCys 2304
Db 241 CAAGGTTCAAGTTTCTGTAAATGAGGCTTATGAGTTGTTGGTGGACAGTCTTGGACATG 300
QY 2304 sGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAlaLysCysPr 2324
Db 301 TCAGAAATCTGGCAATGGAATGAAGATCAAAATCCAAAGTGCATGCTCCCAAGTGCC 360

```

```

QY 2324 oGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyVa 2344
Db 361 AGAGCCGCCCTCTCTGGAAACCCAGCTAGTATTAAAGGAGTTGACCCAGGAGTAGGAGT 420
QY 2344 lValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLe 2364
Db 421 TGTGCACATTTTCTCTGTAAGAGAGGCGATCTCTGCAAGGCCCTCTGTCTCTGTAATGCTT 480
QY 2364 uProSerGlnGlnTrpAsnAspSerPheProValCysLysValLeuLysCysThrProPr 2384
Db 481 GCCATCCAGCAATAGGAATGACTCTTTCCCTGTTTGAAGATTGTTCTTTGTACCCACC 540
QY 2384 oProLeuLysSerPheGlyValProLysProSerSerAlaLeuHisPheGlySerThrVa 2404
Db 541 TCCCTTAATTTCTTGTGTGTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 2404 lLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnPr 2424
Db 601 CAAGTATTTCTGTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 660
QY 2424 oAspGlyThrTrpSerSerProLeuProGluCysValProValGluCysProGlnProGl 2444
Db 661 TGATGGCACCCTGGAGCTCTCCACTGCCAGAAATGTTCCAGTAGAATGTTCCCAACCTGA 720
QY 2444 uGluLysProAsnGlyLysLeuLeuValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTy 2464
Db 721 GGAATCCCAATGGAATCATTGATGTCGAAGGCTTGGCTATCTCAGCACAGCTCTCTA 780
QY 2464 rThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGl 2484
Db 781 TACCTGCAAGCCAGGCTTTGAATTTGGTGGGAAATATCTACCACTCTTGTGGAGAAAATGG 840
QY 2484 yHisTrpLeuGlyGlyLysProThrCysLysAlaLeuGluCysLeuLysProLysGluL 2504
Db 841 TCAGTGGCTTGGAGGAAACCAACATGTAAGCCATTTGAGTGGCTTGAACCAAGAGAT 900
QY 2504 eLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCy 2524
Db 901 TTTGAATGGCAATTTCTTACACGGACCTACACTATGGACAGACCGTTACTACTCTTG 960
QY 2524 sAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTr 2544
Db 961 CAACCGAGGCTTTTCGGCTCGAAGGTCCAGTCCCTGACCTGTTTAGACAGAGGTGATG 1020
QY 2544 pAspValAspAlaProSerCysAsnAlaLeuHisCysAsnSerProGlnProLysGluAs 2564
Db 1021 GGATGTAGATGCCCCATCTTGCATGCAATGCCATCCACTGTGATTTCCCAACCACTTGA 1080
QY 2564 nGlyPheValGluGlyAlaAspTyrSerTyrGlyValAlaLeuLysLeuTyrSerCysPhePr 2584
Db 1081 TGGTTTTGTAGAGGTGCAGATTACAGTATGGTGCCATAATCATCTACAGTTGCTTCCC 1140
QY 2584 oGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerSe 2604
Db 1141 TGGGTTTTCAGGTGGCTGTCATGTCATGCAGACCTGTGGAAGAGTCAGGATGGTCAAGTTC 1200
QY 2604 rLysProThrCysMetProLysAspCysGlyLeuProProHisIleAspPheGlyAspCy 2624
Db 1201 CATCCCAACATGATGCAATAGACTGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY 2624 sThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspMetMetGluValPr 2644
Db 1261 TACTAAACTCAAAGATGACCCAGGATATTTTGGACCAAGAGACGACATGATGGAAGTTCC 1320
QY 2644 oTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsnTh 2664
Db 1321 ATACGTGACTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380
QY 2664 rLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyrTh 2684
Db 1381 AAAGGAGTCTCTCTGCTACACATTCATCAAACTTCTGTATGTTGTTGTTGTTGTTGTTGTT 1440
QY 2684 rCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLysCysGlnGluAspGlyTh 2704

```

1441 CTGTAATCAGGATATGAACCTTCTGGGAACCTCTGCTGATCTGCCGAAGATGGAAAC 1500
2704 rTTPAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaProG1 2724
1501 TTGGAAATGGCAGTGCACCACTCTGCAATTCATTTGAATGTGACTTGCTTACTGCTCTGA 1560
2724 uAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCysLy 2744
1561 AAATGGCTTTTTCGCTTTTACAGAGACTAGCACTAGCAATGGGAAGTGTGTGCAGTATAGCTGTAA 1620
2744 sProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSe 2764
1621 ACCTGGACACATCTAGCAGGCTCTGCACTTAAGGCTTTGTCTAGAGATAGAAATGGAG 1680
2764 rGlyAlaSerProArgCysGluAlaIleSerCysLysIleSerProAsnProValMetAsnG1 2784
1681 TGGTGCTCCCTCCACCGCTGTGAAGCAATTCATGTCAAAAGCCAAATCCAGTCATGAATGG 1740
2784 ySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProG1 2804
1741 ATCCATCAAGGAAGCAACTACACATACCTGAGCACTGTGTACTATGAGTGTGACCCCGG 1800
2804 yTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrpAspGluAs 2824
1801 ATATGTGCTGAATGGCACTGAGAGGAACAATGCCAGGATGACAAATCTGGATGAGGA 1860
2824 pGluProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyGlnVa 2844
1861 TGAGGCCATTTGCATTCCTGTGAGCTGCAGTTCACCCCGCTCAGCCAAATGGCCAGGT 1920
2844 lArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLe 2864
1921 GAGAGGACGAGTACACATTCCTCAAAAGAGATTTGAATACCTTGCAATGAAGGGTTCTT 1980
2864 uLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrPr 2884
1981 GCTTGAGGAGCAGAGTCTGGGTTTGCTTCCCAATGAGAGTTGAGGTGGAGCCACTCC 2040
2884 oAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGlyValThrGluG1 2904
2041 CGACTGTGTGCTGTGCAGATGTGCCACCCGCCACCACTGGCCAAATGGGGTGACGGAGG 2100
2904 yLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHi 2924
2101 CCTGGACTATGGCTTCATGAAGGAAGTAAATTCACATGTCTCAGAGGGCTACATCTGCCA 2160
2924 sGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCy 2944
2161 CGGTGCTCCAAATCACTACCTGTCTGATGATGGCAACTGGGATGACAGATTTCTCTCTG 2220
2944 sLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSe 2964
2221 TAAACCACTCACTGTGAGCACTCTGCAAGATCTTGCCCACTGGTTTCCCTTAATGGTTTTTC 2280
2964 rPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAs 2984
2281 CTTTATTATCATGGGGCCATATACATATCAGTATCAGTGTCTTCTCTGTTATTAAGCTCCATGGAAA 2340
2984 nSerSerArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuPr 3004
2341 TTCATCAAGAGAGTGTCTCTCAATGGCTCTTGAGTGGCAGCTACCTTCTCTGCTGCC 2400
3004 oCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCy 3024
2401 TTGCAGATGTTCCACACCACTAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2460
3024 sGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluTrl 3044
2461 TGGAAAGGCGAGCCCGGATTCAGTGTCTTCAAGGCTTCAAGGCTTCTAGGACTTCTGAAAT 2520
3044 eThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCy 3064

2521 CACCTGTGAAGCCGATGGCCAGTGGAGCTCTGGGTTCCCCACTGTGAACACACTCTTCTTG 2580
3064 sGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGluAs 3084
2581 TGGTTCTCTTCCAAATGATCAAAATGCGTTTCATCAGTGAGACAGCTCTTTGGGAAGAAAA 2640
3084 nValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCy 3104
2641 TGTGATTAATCTTAACGCTGCGAGGTCTGGATATGTATACAAAGGAGTTCAGATCTGATTTG 2700
3104 sThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySe 3124
2701 TACAGAGAAAGGGTATGGAGCCAGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTC 2760
3124 rProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluVa 3144
2761 CCCACGCTGTGCGCAATGSCAGTGGCAACTGGAGAGGCCACCCACCTATGAAAGTGAAGT 2820
3144 lLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPheThrCysG1 3164
2821 GAACTCAGATGTCTGGAAAGTTATACGATGGATACAGATACAGATACATTGACCTGTCA 2880
3164 nLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeuPr 3184
2881 GAAAGATGCTCGTGGTTCCCTGAGAGAAATCTCTGCAAGTCTCTAAATAATGTCTCTCCC 2940
3184 oGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgIleValSe 3204
2941 CGAAATCAATACACATATATCTTTCATGGGACGATTTTCAGTGTGAATGAGGCAAGTTTC 3000
3204 rValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAs 3224
3001 TGTGTCTATGTGAGAGGGTATACCTTTGAGGAGGTTTAAATATCAGTATGTCAGCTTGA 3060
3224 pGlyThrTrpGluProProPheSerAspGluSerCysSerProValSerCysGlyLysPr 3244
3061 TGGAACTGGGAGCCACCACTCTCCGATGAATCTCTGCAAGTCCAGTCTTCTGTGGGAACC 3120
3244 oGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGluSerThrIle11 3264
3121 TGAAGTCCCAAGCAATGGATTTGTGGTGGCAGTAAATACACCTTTTGAAGACCAATAT 3180
3264 eTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAs 3284
3181 TTATCAGTGTGAGCTCTGGCTATGAACTAGAGGGGAACAGGAACGTCTCTGCCAGGAAA 3240
3284 nArgGlnTrpSerGlyValAlaIleCysLysGluThrArgCysGluThrProLeuG1 3304
3241 CAGACAGTGGAGTGGAGGGGTGGCAATATGCAAGAGAGACAGGTGTGAAACTCCACTTGA 3300
3304 uPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSe 3324
3301 ATTTCTCAATGGGAAGCTGACATTTGAATAACAGGACGACTGGAGCCCAACGTGTATATTC 3360
3324 rCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyTh 3344
3361 CTGCAACAGAGCTACAGTCTTGAAGGGCCATCTGAGGCACACTGACAGAAATGGAAC 3420
3344 rTrpSerHisProValProLeuCysLysProAsnProCysProValProPheValIlePr 3364
3421 CTGGAGCCACCCAGTCCCTCTCTGCAAAACCAATCCATGCTCTCTTTTGTGTGATTC 3480
3364 oGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCy 3384
3481 CGAGAAATGCTCTGCTGTCTGAAAGAGGTTTTATGTGTGATCAGATGTGTGTCATCAATG 3540
3384 sArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTr 3404
3541 TAGGAGAGGTTTTCTGCTGACGGGCCACCGCATCATTACTGCAACCCCGCAGAGCTG 3600
3404 pThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAs 3424
3601 GACACAGACAGCGCCAAATGTGAAAAAATCTCATGTGTGCCACAGCTCAGCTCAGTAGAAA 3660

QY 3424 nAlaIleAlaArgGlyValHisThrGlyAspMetIleThrTyrSerCysThrSe 3444
DB 3661 TGCATTCCTCGAGCGTACATTAATATGAGACATGATCACCTACTCATGTATACG 3720
QY 3444 rGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAenGlyThrTrpThrSe 3464
DB 3721 TGGATACATGTTGGAGGGTTCTCGAGGAGTGTGTTTGTAGAAATGGAACATGGACATC 3780
QY 3464 rProProlleCysArgAlaValCysArgPheProCysGlnAenGlyGlyIleCysGlnAr 3484
DB 3781 ACCTCCTATTGACAGAGCTGTCTGCGATTTCCATGTCCAGAAATGGGGCATCTGCCAACG 3840
QY 3484 gProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuGluGluProlleCy 3504
DB 3841 CCCAAATGCTTGTTCCTGTCA-GAGGGCTGGATGGGGCGCTCTGTGAAGAACCAATCTG 3899
QY 3504 sIleLeuProCysLeuAenGlyGlyArgCysValAlaProTyrGlnCysAspCysProPr 3524
DB 3900 CATTCCTCCCTGTCTGAAGAGAGGTGCTGTGTGGCCCTTACCAGTGTGACTGCCCCGCC 3959
QY 3524 oGlyTrpThrGlySerArgCysHisThrAla 3534
DB 3960 TGGCTGGACGGGCTCTGCTGTCTATACAGCT 3990
RESULT 14
US-09-764-853-214
; Sequence 214, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-214
Alignment Scores:
Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 9 Gaps: 0
US-09-977-053-4 (1-3571) x US-09-764-853-214 (1-3804)
QY 2436 ValProValGluCysProGlnProGluGluIleProAenGlyIleIleaspValGlnGly 2455
DB 11 GTTCCAGTAGAATGTCCTCCCACTCAGAAATCCCAATGGAAATCATTTGATGTGACGCC 70
QY 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAen 2475
DB 71 CTTCGCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGCTTTGAAATGGTGGGAAT 130
QY 2476 ThrThrThrLeuCysGlyGluAenGlyHisThrLeuGlyLysProThrCysLysAla 2495
DB 131 ACTACACCCCTTTGTGGAGAAATGGTCTACTGCTTGGAGGAAATCCCAACATGTAAAGCC 190
QY 2496 IleGluCysLeuLysProLysGluIleLeuAenGlyLysPheSerTyrThrAenLeuHis 2515
DB 191 ATTGAGTGCTCGAAACCCAGAGAGATTTTGAATGGCAATTTCTTACACGACCTACAC 250
QY 2516 TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla 2535
DB 251 TATGGACAGACCGTTACCTACTCTTTCGAACCGAGGCTTTGGCTTCGAGGTCCAGTGCC 310

QY 2536 LeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAenAlaIleHis 2555
DB 311 TTGACTCTTTAGACAGAGTGTGGATGTAGATGCCCAATCTTGCATATGCCATCCAC 370
QY 2556 CysAspSerProGlnProIleGluAenGlyPheValGluGlyAlaAspTyrSerTyrGly 2575
DB 371 TGTGATTTCCCAACACCATTCGAAATGGTTTTGTAGAGGTGCAGATTACAGCTATGGT 430
QY 2576 AlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
DB 431 GCCATAATCATCTACAGTTGCTTCCCTGGGTTTCAGGTGGCTGGTCATGCCATGCCAGAC 490
QY 2596 CysGluGluSerGlyTrpSerSerIleProThrCysMetProIleAspCysGlyLeu 2615
DB 491 TGTGAAGAGTCAGGATGTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCCTC 550
QY 2616 ProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGlu 2635
DB 551 CCTCCTCATATAGATTTTGGAGACTGTACTAAACTCAAAGATGACCAGGATATATTTGAG 610
QY 2636 GlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHisLeuGly 2655
DB 611 CAAGAAGACGACATCATGGAAGTTCCATATGTAGTACTCTCTCACCTCTTATCATATTTGGGA 670
QY 2656 AlaValAlaLysThrTrpGluAenThrLysGluSerProAlaThrHisSerSerAenPhe 2675
DB 671 GCAGTGGCTTAAACCTGGGAAATACAAAGGAGTCTCTCTGTCTACACATTCATCAACTTT 730
QY 2676 LeuTyrGlyThrMetValSerTyrThrCysAenProGlyTyrGluLeuLeuGlyAenPro 2695
DB 731 CTGTATGTTACCATGTTTTCATACACCTGTATATCCAGGATATGAACTTCTGGGGAACCT 790
QY 2696 ValLeuIleCysGlnGluAspGlyThrTrpAenGlySerAlaProSerCysIleSerIle 2715
DB 791 GTGCTGATCTGCCAGAAAGATGGAATCTTGGAAATGGCAGTGCACCATCTCTGCAATTTCAAT 850
QY 2716 GluCysAspLeuProThrAlaProGluAenGlyPheLeuArgPheThrGluThrSerMet 2735
DB 851 GAATGTGACTTGCCTACTCTCTGAAATGGCTTTTGGTTTTTACAGAGACTAGCATG 910
QY 2736 GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg 2755
DB 911 GGAAGTGTGTCAGTATAGCTGTAAACCTGGACACATTTCTAGCAGGCTCTGACTTAAAG 970
QY 2756 LeuCysLeuGluAenArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys 2775
DB 971 CTTTGTCTAGAGATAGAAAGTGGAGTGGTGGCTCCCAAGCTGTGAAGCCATTTTCATGC 1030
QY 2776 LysLysProAenProValMetAenGlySerIleLysGlySerAenTyrThrTyrLeuSer 2795
DB 1031 AAAAAGCCAAATCCAGTCATGATGATGATCCATCAAGGAGCAACTACACATACCTGAGC 1090
QY 2796 ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAenGlyThrGluArgArgThrCys 2815
DB 1091 AGTTTGTACTAGTGTGACCCCGATATGTGTGATGTGCACTGAGAGGAGAAACATGC 1150
QY 2816 GlnAspAspLysAenTrpAspGluAspGluProlleCysIleProValAspCysSerSer 2835
DB 1151 CAGGATGCAAAATCTGGGATGAGATGAGCCCATTTGATTCCTGTGGACTGCAGTTCA 1210
QY 2836 ProProValSerAlaAenGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle 2855
DB 1211 CCCCCAGTCTCAGCCCATGCCCCAGGTGAGAGGAGCAGTACACATTCCAAAGAGATTT 1270
QY 2856 GluTyrThrCysAenGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla 2875
DB 1271 GAATACACTTGCATTAAGGGTTCTTGTCTGAGGAGCCAGGAGTCGGGTTTGTCTTGCC 1330
QY 2876 AenGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895
DB 1331 AATGGAAGTTGGAGTGGAGCCACTCCCGACTGTGTGCTGTGTAGATGTGCCACCCCGCCA 1390

2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe 2915
1391 CAATGGCAATGGGGTACCGAAGGCTTGACTATGGCTTCATGAGGAAGTAACATTC 1450
2916 HisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly 2935
1451 CACTGTCAAGAGGGCTACATCTTGCACGGTCTCCANAACCTCACCTGTCACTAGATGGC 1510
2936 AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeu 2955
1511 AACTGGGATGAGAGATTCCTCTGTAAACAGCTCAACTGTGACCTCCCTGAAAGATCTT 1570
2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys 2975
1571 GCCCATGGTTCCCTTAATGGTTTTCCTTATTCATGGGGGCCATATACAGATACAGTGC 1630
2976 PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp 2995
1631 TTTCTGGTTAATAGCTCCATGGAAATTCATCAAGAGGTGCCTCTCCAAATGGCTCTCTGG 1690
2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
1691 AGTGGCAGCTCACCTTCCTGCTGCTTGCAGATGTTCCACACCAAGTAATTAATATATGA 1750
3016 ThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly 3035
1751 ACTGTCAATGGACAGATTTTGACTGTGGAAGAGGACGCCCGGATTCAGTGTTCAGAGGC 1810
3036 PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly 3055
1811 TTCAAGCTCCTAGGACTTCTGAAATCACCTGTGAAGCGCATGGCCAGTGGAGCTCTGGG 1870
3056 PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
1871 TTCCCCCATCTGTGAACACATCTTGTGGTTCCTCTCCAAATGATACCAATGCGTTCATC 1930
3076 SerGluThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
1931 AGTCAGACCACTCTTGGAGAGAAATGTGATTACTACACTGCACTGCACTGTGGATATGTC 1990
3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpPro 3115
1991 ATACAAGGCAGTTCAGATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCA 2050
3116 ValCysGluProLeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGly 3135
2051 GTCTGTGAGCCCTTGTCTGTGGGTCCCACCGTCTGTGCCAATGCAAGTGGCAACTGGA 2110
3136 GluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
2111 GAGGCACACACTATGAAGTGAAGTGAAATCTCAGATGCTCTGGAAGGTTATACGATGGAT 2170
3156 ThrAspThrAspThrCysGlnLysAspGlyArgTrpPheProGluArgIleSer 3175
2171 ACAGATACAGATACATTCACCTGTGAGAAAGATGTCGTCTGCTTCCCTGAGAGATCTCC 2230
3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
2231 TGCAGTCTCTAAATAATGCTCTCCCGGAAACATAACACATATACCTTGTACATGGGGAC 2290
3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
2291 GATTTCAATGATAGCAAGTTCTGTGTGTCATGTGCAAGAGGTATACCTTTGAGGGA 2350
3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer 3235
2351 GTTAACATATCAGTATGTGAGTTCATGTAACCTGGGAGCCACCAATCTCCGATGAATCT 2410
3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer 3255
2411 TGCAGTCCAGTTCTTGTGGGAAACCTGAAAGTCCAGAAACATGGAATTTGTGGTGGCAGT 2470
3256 LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGly 3275

2471 AAATACACCTTTGAAAGCACCAATTTATTTATCAGTGTGAGCCCTGGCTATGAACCTAGAGGG 2530
3276 AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLys 3295
2531 AACAGGGAACCGTGTCTGCCAGGAGACACAGACTGTGGAGGGGTGGCAATATGCAAA 2590
3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg 3315
2591 GAGACAGGTGTGAAATCTCACTTGAATTTCTCAATGGGAAGCTGACATTTGAAACAGG 2650
3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
2651 ACGACTGGACCAACAGTGTATATTCCTGCAACAGAGCTACAGTCTTGAAGGCCCATCT 2710
3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn 3355
2711 GAGGCACACTGCACAGAAATGGAACTGGAGGCCACCCAGTCCCTCTCTGCAAAACCAAT 2770
3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr 3375
2771 CCATGCCCTGTCTCTTGTGATTCCTGAGATGCTCTGCTCTGAAAAAGAGTTTAT 2830
3376 ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
2831 GTTGATCAGAAATGTCCATCAATGTAGGAAGGTTTCTGCTGACAGGCCACCGCATC 2890
3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSer 3415
2891 ATTACCTGCAACCCCGACGACCTGGACACAGACAGCGCAATGTGAAAAATCTCA 2950
3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
2951 TGTGTTCACACAGCTCAGTAGAAAAATCAATTTGTCAGAGGCTACATTTATCAATATGA 3010
3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
3011 GACATGATCACCTACTCATGTTACAGTGCATACATGTTGGAGGGTTTCTGAGGAGTGT 3070
3456 CysLeuGluAsnGlyThrTrpThrSerProPheCysArgAlaValCysArgPhePro 3475
3071 TGTTAGAAAAATGGAACATGGACATCACCTCTCTATTGTGAGAGCTGTCTGCAATTCCA 3130
3476 CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMet 3495
3131 TGTCAATGGGGCATCTGCCACGCCCAATGCTGTCTGTCAGAGGGCTGATG 3190
3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
3191 GGGCGCTCTGTGAAGAACCAATCTGCATTTCTCTCTGTCGACGAGGCTCGCTGTGTG 3250
3516 AlaProTyrGlnCysAspCysProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
3251 GCCCTTACCAAGTGTACTGCCCGCTGCTGGAGCGGGGTCTCGCTGTCTATACAGCTGT 3310
3536 CysGlnSerProCysLeuAsnGlyCysValArgCysValArgProAsnArgCysHisCysLeu 3555
3311 TGCCAGTCTCTCTCTCTTAAATGGTGAATGTGTGAACCAACCAACCGATGTCACTGTCTT 3370
3556 SerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe 3571
3371 TCTTCTGGAGGGACATAACTTTCCAGGAAAAAGAGGAGGAGCTGGGTTT 3418

RESULT 15

US-09-764-898-60

; Sequence 60, Application US/09764898

; Patent No. US20020090673A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PU201

; CURRENT APPLICATION NUMBER: US/09/764,898

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 311

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 60

; LENGTH: 3804

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-898-60

Alignment Scores:

Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-764-898-60 (1-3804)

QY	2436	valprovalglnCysProGlnProGluGluLeuProAsnGlyLeuLeuAspValGlnGly	2455
DB	11	GTTCAGTAGAATGTCCCAACCTCGAATAATCCCAATGGAATCATTTGATGTCAGGC	70
QY	2456	LeuAlaTyrLeuSerThrAlaLeuTyrThrCysValysProGlyPheGluLeuValGlyAsn	2475
DB	71	CTTGCCCTATCTCAGCAGCTCTCTATACCTGCAAGCCAGCTTTGAATTTGGGGAAAT	130
QY	2476	ThrThrThrLeuCysGlyGluAsnGlyHisTyrLeuGlyGlyPheThrCysLysAla	2495
DB	131	ACTACCACCTTTGTGAGAAATGGTCACTGGCTTGAGGAAACCAACATGTAAAGCC	190
QY	2496	IleGluCysLeuLysProLysGluLeuAsnGlyLysPheSerThrThrAspLeuHis	2515
DB	191	ATTGAGTGCCTGAAACCCCAAGAGATTTTGAATGGCAATTTCTTTACACGACCTACAC	250
QY	2516	TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla	2535
DB	251	TATGGACAGACCTTACCTACTCTTGACACGAGGCTTTGGCTCGAGAGTCCAGTGCC	310
QY	2536	LeuThrCysLeuGluThrGlyAspThrAspValAspAlaProSerCysAsnAlaIleHis	2555
DB	311	TTGACCTGTTTACAGACAGGTGATGGGATGGATGGATGGATGGATGGATGGATGGAT	370
QY	2556	CysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGly	2575
DB	371	TGTGATTTCCCAACACCCATTTGAAATGGTTTGTAGAGGTGACAGATTACAGCTATGGT	430
QY	2576	AlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr	2595
DB	431	GCATATATCATCTACAGTTGCTTCTTCTGGGTTTCAGGTGGTGGTGGTGGTGGTGGT	490
QY	2596	CysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeu	2615
DB	491	TGTGAAGAGTCAGAGTGTCAAGTTCCATCCCAACATGATGTCATAGACTGTGGCTC	550
QY	2616	ProProHisIleAspPheGlyAspCysThrLysLeuLysAsnAspGlnGlyTyrPheGlu	2635
DB	551	CTCTCCCTCATATAGATTTTGGAGACTGTACTAAACTCAAGATGACCAAGGATATTTGAG	610
QY	2636	GlnGluAspAspMetMetGluValProTyrValThrProHisProTyrHisLeuGly	2655
DB	611	CAAGAAGACGACATGATGGAGTTCCATATGATGACTCTCCATCCCTCTTATCATTTGGGA	670
QY	2656	AlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhe	2675
DB	671	GCAGTGGCTAAAACCTGGGAAATAACAAGAGTCTCTCTGCTACACATTCATCAAACTTT	730
QY	2676	LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuGlyAsnPro	2695
DB	731	CTGTATGGTACCATGTTTTCATACACTGTGTAATCCAGGATATGAACTTCTGGGGAACCT	790
QY	2696	ValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle	2715

DB	791	GTGCTGATCTGCCAGGAAGTGGAACTTGGAAATGGCAGTCACCATCTCTGCATTTCATTT	850
QY	2716	GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMet	2735
DB	851	GAATGTGACTTGCCTACTGCTCTCTGAAATGGCTTTTGGCGTTTTCACAGAGACTAGCATG	910
QY	2736	GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg	2755
DB	911	GGAAGTCTGTGCAGTATAGCTGTAAACCTGGACACATCTAGCAGGCTCTGACTTAAGG	970
QY	2756	LeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys	2775
DB	971	CTTTGCTTAGAATAGAAAGTGGAGTGGCTCTCCCAACCTGTGAAGCCATTTTCATGC	1030
QY	2776	LysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSer	2795
DB	1031	AAAAAGCCAAATCCAGTCATGATGGATCCATCAAGGAAGCACTACACATACCTGAGC	1090
QY	2796	ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCys	2815
DB	1091	ACGTTGTACTATGAGTGTGACCCCGGATATGTCTGAATGGCAGTACAGAGGAAACATGC	1150
QY	2816	GlnAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer	2835
DB	1151	CAGATGACAAACCTGGATGAGGATGAGCCATTTGCATTTCTGTGAGCTGACAGTTCA	1210
QY	2836	ProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle	2855
DB	1211	CCCCCAGTCTCAGCCAAATGGCCAGGTGAGAGGAGACGAGTACACATTTCCAAAAGAGATT	1270
QY	2856	GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla	2875
DB	1271	GAATACACTTGCATGAAGGGTCTTCTGCTGAGGAGCCAGAGTCCGGGTTTGTCTGCC	1330
QY	2876	AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
DB	1331	AATGGAAGTTGGAGTGGAGCCACTCCCGAGCTGTGTGCTGTGCAGATGTGCCACCCGCCA	1390
QY	2896	GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe	2915
DB	1391	CAACTGCCCAATGGGTGACGAGGCTTGCAGTATGGCTTTCATGAAGGAAGTAAATTC	1450
QY	2916	HisCysHisGluGlyTyrIleLeuHisGlyValaProLysLeuThrCysGlnSerAspGly	2935
DB	1451	CATGTCCAGGAGGTACATCTTGACGGTGTCTCCAAACCTCACCTGTGCTCAGTACAGTGC	1510
QY	2936	AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeu	2955
DB	1511	AATGGGATGACAGAGATTCTCTCTGTAAACCACTCACTGTGGACCTCTCTGAAGATCTT	1570
QY	2956	AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys	2975
DB	1571	GCCCATGGTTTCCCTAATGGTTTCTTCTTATTCATGGGGGCCATATACAGTATCAGTGC	1630
QY	2976	PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp	2995
DB	1631	TTTCTGGTTTATAGCTCCATGGAATTCATCAAGAAGGTGCTCTCCAAATGGCTCTCGG	1690
QY	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly	3015
DB	1691	AGTGGCAGCTCACCTTCTCTGCTGCTTGCAGATGTTCCACACCAAGTAATTAATATGGA	1750
QY	3016	ThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly	3035
DB	1751	ACTGTCAAATGGGACAGATTTTGTGTAAGAGGAGCCCGGATTCAGTGTCTTCAAGAGC	1810
QY	3036	PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly	3055
DB	1811	TTACAGCTCTAGACTTTCTGAAATCACCTGTGAGCCGATGCGGAGTGGGCTCTGGG	1870
QY	3056	PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle	3075
DB	1871	TTCCCCCAGCTGTGAACACACTTCTTGTGGTTCTCTTCCAAATGATACCAATGCGTTCATC	1930

```
QY 3076 SerGluThrSerSerSerTyrIleValIleThrThrSerCysArgSerGlyTyrVal 3095
DB 1931 AGTGAACAGCCTCTTGTGAAGGAAAATGTGATAACTTACAGCTGCAGGCTCGATATGTC 1990
QY 3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTyrSerGlnProTyrPro 3115
DB 1991 ATACAGGAGGTTGAGATCTGATTTGTACAGAGAAAGGGGTATGGAGCAGGCTTATCCA 2050
QY 3116 ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly 3135
DB 2051 GTCTGTGAGCCCTTGTCTCTGTGGTCCCGACCGTCTGTGCGCAATGTCAGTGGCAACTGGA 2110
QY 3136 GluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
DB 2111 GAGGCACACACCTATGAAGAAGTGAAGTGAACATCAGATGTCAGAAAGGTTATACGATGGAT 2170
QY 3156 ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTyrPheProGluArgIleSer 3175
DB 2171 ACAGATACAGATACATTCACCTGTGAGAAAGATGCTGCTGCTTCCCTGAGAGATCTCC 2230
QY 3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
DB 2231 TGCAGTCTCTAAAAATGCTCTCTCCGGAACACATAACACATATATCTTGTATCATGGGAC 2290
QY 3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
DB 2291 GATTTTCAGTGTGAATAGGCAAGTTTCTGTGTCTATGTGCAAGAGGATATACCTTTTCAGGGA 2350
QY 3216 ValAsnIleSerValCysGlnLeuAspGlyThrTyrGluProPheSerAspGluSer 3235
DB 2351 GTTAAACATATCATGATGTCAGCTTCAATGAGCACTGGAGCCACCAATTTCTCCATGATCT 2410
QY 3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer 3255
DB 2411 TGCAGTCCAGTTTCTGTGGGAACCTGAAAGTCCAGAACATGATTTGTGGTGGCAGT 2470
QY 3256 LysTyrThrPheGluSerThrIleLysGlnCysGluProGlyTyrGluLeuGluGly 3275
DB 2471 AAATACACCTTTTGAAGCACCAATTTATTCATGCTGAGCCTGGCTGTATGAACATAGAGGG 2530
QY 3276 AsnArgGluArgValCysGlnGluAsnArgGlnTyrSerGlyValAlaIleCysLys 3295
DB 2531 AACAGGAAAGTGTCTGCAGAGAGAACAGACAGTGGAGTGGAGGGTGGCATATGCAAA 2590
QY 3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg 3315
DB 2591 GAGACCAAGTGTGAAACTCCACTTGAATTTCTCAATGGGAAAGCTGACATTCGCAAAACAGG 2650
QY 3316 ThrThrGlyProAsnValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
DB 2651 ACGACTGGACCCCAACGTGGTATATTCCTGCAACAGAGGCTACAGTCTTTGAAGGGCCATCT 2710
QY 3336 GluAlaHisCysThrGluAsnGlyThrTyrPheHisProValProLeuCysLysProAsn 3355
DB 2711 GAGGCACACTGCAGAGAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGCAAAACCAAT 2770
QY 3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr 3375
DB 2771 CCATGCCCTGTCTCTTTTGTGATTCCTGAGAAATGTTCTGCTGCTGAGGGCCACCGCATC 2830
QY 3376 ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGluGlyHisGlyIle 3395
DB 2831 GTTGTATCAGATGTGTGCATCAAAATGTAGGGGAAGGTTTCTGCTGAGGGCCACCGCATC 2890
QY 3396 IleThrCysAsnProAspGluThrTyrThrGlnThrSerAlaLysCysGluLysIleSer 3415
DB 2891 ATTACTGCACCCCGAGAGAGCTGGACACAGACAGAGCCCAATGTGAAAAATCTCA 2950
QY 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
DB 2951 TGTGTGCCACAGCTCAGGTAGAAAATGCAATTTGCTCGAGGGCTATCATTTATCAATATGGA 3010
```

```
QY 3436 AspMetIleThrThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
DB 3011 GACATGATCACCTACTCTCATGTTACAGTGGATACATGTTGGAGGGTTTCTTGAGAGTGT 3070
QY 3456 CysLeuGluAsnGlyThrTyrThrSerProProIleCysArgAlaValCysArgPhePro 3475
DB 3071 TGTTTAGAAAATGGAACATGGACATCACCTCTCTATTTGAGAGCTGTCTGTGATTTCCA 3130
QY 3476 CysGlnAsnGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMet 3495
DB 3131 TGTTCAGAAATGGGGCATCTGCCACAGCCCAAAATGCTTGTCTCTCCAGAGGGCTGGAG 3190
QY 3496 GlyArgLeuLeuGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
DB 3191 GGGCCCTCTGTGAGAAACCAATCTGCATTTCTCTCTGCTGAACGAGAGTCTGCTGTG 3250
QY 3516 AlaProTyrGlnCysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
DB 3251 GCGCTTACCACTGTGACTGCCCGCTGCTGGACGGGGTCTCGCTGTCTATACAGCTGT 3310
QY 3536 CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu 3555
DB 3311 TGCCAGTCTCCCTGCTTAAATGGTGGAAATGTGTAAAGACCAACCGATGTCTGTCT 3370
QY 3556 SerSerTyrThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe 3571
DB 3371 TCTTCTTGACGGGACATACTGTTCCAGGAAAGAGGAGACTGGGT 3418

RESULT 16
US-09-764-881-21
; Sequence 21, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-881-21

Alignment Scores:
Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 10 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-764-881-21 (1-3804)
QY 2436 ValProValGluCysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGly 2455
DB 11 GTTCCAGTAGAATGTCCCACTGAGAAAATCCCCAATGGAAATCAATTGATGTGCAAGGC 70
QY 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsn 2475
DB 71 CTGTCTATCTCAGCACAGCTCTCTATACCTGCAGCCAGGCTTTGAATTTGGTGGGAAT 130
QY 2476 ThrThrThrLeuCysGlyGluAsnGlyHisIleTyrLeuGlyLysProThrCysLysAla 2495
DB 131 ACTACCAACCTTTGTGGAGAAAATGCTCACTGGCTTGGAGGAAACCAACATGTAAAGCC 190
QY 2496 IleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHis 2515
DB 191 ATTGAGTGCCTGAAACCCCAAGGAGATTTTGAATGGCAATTTCTTTACACGGGACCTACAC 250
```

QY 2516 TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla 2535
 DB 251 TATGGACAGACCGTTACCTACTCTTGACACCGAGGCTTTTCGGCTCGAAGGTCCCGAGTGC 310
 QY 2536 LeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHis 2555
 DB 311 TTGACCTGTTTACAGACAGGTGATGGATGTAGATGCCCACTCTTGCAGTCCATCCAC 370
 QY 2556 CysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGly 2575
 DB 371 TGTGATTTCCCAACAACCACTTGAATGTTTGTAGAGGTGAGATACAGATTACAGATTGTT 430
 QY 2576 AlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
 DB 431 GCCATATCATCATACAGTTCCTCCCTGGGTTCAGGTGGCTGATCCATGCAGAC 490
 QY 2596 CysGluGluSerGlyTrpSerSerIleProThrCysMetProIleAspCysGlyLeu 2615
 DB 491 TGTGAGAGTCAGATGCTCAGTTCCATCCCAACATGTATGCCATAGACTTGGGCTC 550
 QY 2616 ProProHisIleAspPheGlyAspCysThrLysLeuLysAspGlnGlyTyrPheGlu 2635
 DB 551 CCTCCTCATATAGATTTGGAGACTGTACTAACTCAAGATGACCAAGGATATTTTGG 610
 QY 2636 GlnGluAspAspMetGluValProTyrValThrProHisProProTyrHisLeuGly 2655
 DB 611 CAAGAAGACACATGATGGAAGTTCCATATGTGACTCCTCACCTCTTATCATTTTGGGA 670
 QY 2656 AlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhe 2675
 DB 671 CGAGTGGCTAAACCTGGGAATAATACAAAGAGTCTCTCTGACTACATTCATCAACTTT 730
 QY 2676 LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuGlyAsnPro 2695
 DB 731 CTGTATGGTACCATGGTTTCATACACCTGTAATCCAGGATATGAATCTCTGGGAACCT 790
 QY 2696 ValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle 2715
 DB 791 GTGCTGATCTGCCAGGAAGATGAGACTTGAATGGCAGTGCCACCATCTCGCATTTCAAT 850
 QY 2716 GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMet 2735
 DB 851 GAATGTGACTTGCTACTGCTCTCGAAATGGCTTTTGGTTTACAGAGACTAGCATG 910
 QY 2736 GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg 2755
 DB 911 GGAAGTGTGTGACGATATAGCTGTAAACCTGGACACATCTAGCAGGCTCTGACTTAAG 970
 QY 2756 LeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys 2775
 DB 971 CTTTGTCTAGAGAAATAGAAAGTGGAGTGGTCCCTCCACGCTGTGAAGCCATTTCAAT 1030
 QY 2776 LysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSer 2795
 DB 1031 AAAAGCCAAATCCAGTCATGATGATGATCCATCAAGAGAGCAACTACATACCTGAGC 1090
 QY 2796 ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCys 2815
 DB 1091 ACCTGTGACTATGAGTGTGACCCCGGATATGTGTGAATGGCACTGAGAGGAGAACATGC 1150
 QY 2816 GlnAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer 2835
 DB 1151 CAGGATGACAAAACCTGGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 1210
 QY 2836 ProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle 2855
 DB 1211 CCCCCAGTCTACCCCAATGGCCAGGTGAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1270
 QY 2856 GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla 2875
 DB 1271 GAATACACTTGAATGAAGGTTCTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1330
 QY 2876 AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895

DB 1331 AATGGAGTTGGAGTGGAGCCACTCCCGACTGTGTGCTGTGCAGATGTGCCACCGCCA 1390
 QY 2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe 2915
 DB 1391 CAATGGCCCAATGGGTGACGAGGCTCTGACTATGGCTTTTCATGAAGAAAGTAACATTC 1450
 QY 2916 HisCysHisGluGlyTyrIleLeuHisGlyValaProLysLeuThrCysGlnSerAspGly 2935
 DB 1451 CACTGTCCAGGGCTACATCTTGACGGTGTCTCAAACTCACCTGTCTGCTCAGTCAGTGGC 1510
 QY 2936 AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeu 2955
 DB 1511 AACTGGATGCAGAGATTCTCTCTGTAAACAGTCAACTGTGGACCTCTCTGAAGATCTT 1570
 QY 2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys 2975
 DB 1571 GCCCATGTTTCCCTAATGTTTCTTATTCATGGGGCCATATACAGTATCAGTGC 1630
 QY 2976 PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp 2995
 DB 1631 TTTCTCTGTTTATAGCTTCCATGGAAATTCATCAAGAAGGTGCTCTTCCATGGCTCTCGG 1690
 QY 2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
 DB 1691 AGTGGACGCTCACTTCTGCTGCTGCTTGCAGATGTTCCACACAGTAATTAATATGA 1750
 QY 3016 ThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly 3035
 DB 1751 ACTGTCAATGGGACAGATTTTGACTGTGGAAGGAGCCCGGATTCAGTGTCTCAAGGC 1810
 QY 3036 PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly 3055
 DB 1811 TTAAGTCTCTAGACATTTCTGAAATCACCTGTGAGCCGATGGCCAGTGTGAGTCTCGG 1870
 QY 3056 PheProHisCysGlnHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
 DB 1871 TTCCCCCCTGTGAACACACACTTCTTGTGTGTTCTCTCCATGATACCAAAATCGTTCATC 1930
 QY 3076 SerGluThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
 DB 1931 AGTGAGACCACTCTCTGGAAGGAAATGTGATACTTACAGCTGCAGGTCTGGATATGTC 1990
 QY 3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpPro 3115
 DB 1991 ATACAGCGAGTTCAGATCTGATTTGTACAGAAAGGGGTATGGAGCCAGCTTATCCA 2050
 QY 3116 ValCysGluProLeuSerCysGlySerProSerProSerValAlaAsnAlaValAlaThrGly 3135
 DB 2051 GTCTGTGAGCCCTTGTCTGTGGGTCCCACTGTGTGCGCAATGCGATGCGCAACTGA 2110
 QY 3136 GluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
 DB 2111 GAGGCAACACCTATGAAGTGAAGTGAATCTCAGATGTCTGGAAGGTATATACGATGAT 2170
 QY 3156 ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgLysSer 3175
 DB 2171 ACAGATACAGATACATTCACCTGTGCAGAAAGTGTGCTGGTTCCTTCCAGAGAAATCTCC 2230
 QY 3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
 DB 2231 TGCAGTCTTAATAAATGTCTCTCCCGGAAACATACACATATATCTTGTATACATGGGAC 2290
 QY 3196 AspPheSerValAsnArgGlnValSerCysAlaGluGlyTyrThrPheGluGly 3215
 DB 2291 GAATTCAGTGTGAATAGGCAAGTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2350
 QY 3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer 3235
 DB 2351 GTTAACATATCAGTATGTGAGTGTGAGAACCTGGGAGGCCCATTTCTCCGATGATCT 2410
 QY 3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer 3255

```
Db 2411 TGCACTCCAGTTCTTGTGGGAAACCTGAAAGCTCCAGAACATGGATTTTGTGTGGCTG 2470
Qy 3256 LysfYrThrPheGluSerThrIleileYrGlnCysGluProGlyYrGluLeuGluGly 3275
Db 2471 AATATACACTTTGAAGACACATTTATTTATCATGTGTGAGCTGGCTGTATGAACATAGAGGG 2530
Qy 3276 AenArgGluArgValCysGlnGluAenArgGlnTrpSerGlyGlyValAlaIleCysLys 3295
Db 2531 AACAGGACAGCTGTCTGCCAGGAGAACACACAGTGGAGTGGAGGGTGGCAATATGCAAA 2590
Qy 3296 GluThrArgCysGluThrProLeuGluPheLeuAenGlyLysAlaAaspIleGluAenArg 3315
Db 2591 GAGACAGAGTGAAACTCCATCTGAATTTCTCAATGGGAAGCTGACATTTGAACACAGG 2650
Qy 3316 ThrThrGlyProAenValValYrSerCysAenArgGlyYrSerLeuGluGlyProSer 3335
Db 2651 ACGACTGGACCAACGTTGTATATTCTCTCAACAGAGGCTACAGTCTTGAAGGGCCATCT 2710
Qy 3336 GluAlaHisCysThrGluAenGlyThrTrpSerHisProValProLeuCysLysProAen 3355
Db 2711 GAGGCACACTGCACAGAAATGAACCTGGAGCCACCCAGTCCCTCTCTGCAAAACCAAT 2770
Qy 3356 ProCysProValProPheValIleProGluAenAlaLeuSerGluLysGluPheYr 3375
Db 2771 CCATGCCCTTGTTCTTTGTGATTTCCGAGAAATGCTCTGCTGCTGNAAGAGGTTTTAT 2830
Qy 3376 ValAenGluAenValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
Db 2831 GTTGATCAGATGTGTCATCAATCAATGTAGGGAAGGTTTTCTGCTGCAGGGCCACGGCATC 2890
Qy 3396 IleThrCysAenProAenGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSer 3415
Db 2891 ATTACCTGCAACCCGACAGAGCTGGACACAGACAGGCGCCAAATGTGAAAAATCTCA 2950
Qy 3416 CysGlyProProAlaHisValGluAenAlaIleAlaArgGlyValHisYrGlnTrpGly 3435
Db 2951 TGTGGTCCACCACTCAGCTAGGAAATGCAATTCCTCGAGGCGTACATTTATCATATGGA 3010
Qy 3436 AspMetIleThrYrSerCysYrSerGlyYrMetLeuGluGlyPheLeuArgSerVal 3455
Db 3011 GACATGATCACCTACTCATGTTACATGTCGATACATGTTGGAGGGTTTTCTGAGGAGTGT 3070
Qy 3456 CysLeuGluAenGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
Db 3071 TGTATTAGAAATGGAACATGGACATCACTCTCTATTTCAGAGCTGTCTGTGTCGATTTCCA 3130
Qy 3476 CysGlnAenGlyGlyIleCysGlnArgProAenAlaCysSerCysProGluGlyTrpMet 3495
Db 3131 TGTCAAGATGGGGGCACTGCCAACGCCCAATGCTTGTCTCTGTCAGAGGGCTGGATG 3190
Qy 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAenGlyGlyArgCysVal 3515
Db 3191 GGGCGCTCTGTGAGAAACCAATCTGCATTTCTTCCCTGTCTGAACGGAGGTGCTGTGTG 3250
Qy 3516 AlaProYrGlnCysAenCysProProGlyTrpThrGlySerArgCysHisThrAlaVal 3535
Db 3251 GCCCTTACCAAGTGTGACTGCCCGCTGCGTGGACCGGGTCTCGCTGTATACAGCTGT 3310
Qy 3536 CysGlnSerProCysLeuAenGlyGlyLysCysValArgProAenArgCysHisCysLeu 3555
Db 3311 TGCAGTCTCCCTGCTTAAATGGTGGAAATGTGTAAAGACCAACCGATGTCTGTCTT 3370
Qy 3556 SerSerTrpThrGlyHisAenCysSerArgLysArgThrGlyPhe 3571
Db 3371 TCTTCTTGGACGGACATACCTGTTCCAGGAAAGGAGGACTGGGTTT 3418
```

RESULT 17

```
US-09-764-875-275
; Sequence 275, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 275
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-875-275

Alignment Scores:
Score: 0 Length: 3804
Percent Similarity: 6452.00 Matches: 1135
Best Local Similarity: 100.00% Conservative: 1
Query Match: 99.91% Mismatches: 0
DB: 32.30% Indels: 0
Gaps: 0
```

US-09-977-053-4 (1-3571) x US-09-764-875-275 (1-3804)

```
Qy 2436 ValProValGluCysProGlnProGluGluIleProAenGlyIleIleAepValGlnGly 2455
Db 11 GTTCCAGTAGAATGTCCTCCCACTGAGAAATCCCAATGGAATCATTTGATGTGCAAGGC 70
Qy 2456 LeuAlaYrLeuSerThrAlaLeuYrThrCysLysProGlyPheGluLeuValGlyAen 2475
Db 71 CTTCCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGCTTTGAAATTTGGTGGGAAAT 130
Qy 2476 ThrThrThrLeuCysGlyGluAenGlyHisTrpLeuGlyGlyLysProThrCysLysAla 2495
Db 131 ACTACCACCCTTTGTGGAGAAATGTCCTGCTTGGAGGAAACCAACATGTGAAGCC 190
Qy 2496 IleGluCysLeuLysProLysGluIleLeuAenGlyLysPheSerYrThrAepLeuHis 2515
Db 191 ATTGAGTGCCTGAAACCCCAAGGAGATTTTGAATGGCAAAATCTCTTACAGGACCTACAC 250
Qy 2516 TyrGlyGlnThrValThrYrSerCysAenArgGlyPheArgLeuGluGlyProSerAla 2535
Db 251 TATGGACAGACCGCTTACCTACTCTTCGCAACCGAGGCTTTCCGGCTCGAAGGTCCAGTGCC 310
Qy 2536 LeuThrCysLeuGluThrGlyAepTrpAepValAaspAlaProSerCysAenAlaIleHis 2555
Db 311 TTGACCTGTTTAGAGACAGTGATTTGGATGTAGATGCCCATCTTGCATATGCATCCAC 370
Qy 2556 CysAepSerProGlnProIleGluAenGlyPheValGluGlyAlaAepYrSerYrGly 2575
Db 371 TGTGATTTCCCAACCAACCCATTGAAATGGTTTGTAGAGGTGCAGATTACAGCTATGTT 430
Qy 2576 AlaIleIleIleYrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
Db 431 GCCATAATCATCTACAGTTGCTTCCCTGGGTTCAGGTGGCTTGTGTATGCCATAGCATGCAGACC 490
Qy 2596 CysGluGluSerGlyTrpSerSerIleProThrCysMetProIleAepCysGlyLeu 2615
Db 491 TGTGAAGATCAGATGGTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCTC 550
Qy 2616 ProProHisIleAepPheGlyAepCysThrLysLeuLysAepAepGlnGlyYrPheGlu 2635
Db 551 CCTCTCATATAGATTTGGAGACTGTACTAAACTCAAGATGACCCAGGATATTTTGAG 610
Qy 2636 GlnGluAaspAepMetMetGluValProYrValThrProHisProProYrHisLeuGly 2655
Db 611 CAAGAAGACGACATGATGGAAATTCATATGTGACTCTCACCCTCTCTTATCATTTGGGA 670
Qy 2656 AlaValAlaLysThrTrpGluAenThrLysGluSerProAlaThrHisSerSerAenPhe 2675
Db 671 GCAGTGGCTTAAACCTGGGAAATATACAAAGAGGTCTCTCTGTACACATTCATCAAACTTT 730
Qy 2676 LeuYrGlyThrMetValSerYrThrCysAenProGlyYrGluLeuLeuGlyAenPro 2695
Db 731 CTGTATGGTACCATGGTTTTCATACACCTGTAATCCAGGATATGAACCTCTCTGGGAAACCT 790
```

Qy	2696	VallLeuIleCysGlnGluAspGlyThrTrpAasnGlySerAlapProSerCysIleSerile	2715
Dd	791	GTGCTGATCGCCAGGAAGAATTGGAAATGCAGTCACCATCTTGCATTTCAATT	850
Qy	2716	GluCysaspLeuProThralaProGluaSncGlyPheLeuarqPheThrGluThrSerMet	2735
Dd	851	GAATGTGACTTGCCCTACTGCTCTCTGAAAATGGCTTTTTTGGGTTTTTACAGAGACTAGCATG	910
Qy	2736	GlySerAlaValGlnTyrrSerCysLysProGlyHisIleLeualagLySerAspLeuarq	2755
Dd	911	GGAAAGTGTGTGCGAGTAGCTGTAAACCTGGACACATTTCTAGCGAGCTCTGACTTHAAG	970
Qy	2756	LeuCysLeuGluAanArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys	2775
Dd	971	CTTTGTCTAGAGATAGAAAGTGGAGTGGTGCTCCCCACGCTGTGAAGCCAATTTTCATGC	1030
Qy	2776	LysLysProAenProValMetAasnGlySerIleLysGlySerAenTyrrThrTyrrLeuSer	2795
Dd	1031	AAAAAGCCNAATCCAGTTCATGAATGGATGCCATCAAAGGAACAACACTACATACCTCAGC	1090
Qy	2796	ThrLeuTyrrTyrrGluCysAspProGlyTyrrValLeuAenGlyThrGluArqArqThrCys	2815
Dd	1091	ACGTTGTACTATGATGTGACCCCGGATATGTGTGAATGGCACTGAGAGGAGAACATGC	1150
Qy	2816	GlnAspAspLysAasnTrpAspGluAspGluProIleCysIleProValAspCysSerSer	2835
Dd	1151	CAGAAATGACAAANAACCTGGGATGAGATGAGGCCAATTTGCATTCCTGTGACAGTTC	1210
Qy	2836	ProProValSerAlaAenGlyGlnValArgGlyAspGluTyrrThrPheGlnLysGluIle	2855
Dd	1211	CCCCAGTCTCAGCCNAATGGCCAGTGTGAGAGGAGACGAGTACACATTTCCAAAAGAGATT	1270
Qy	2856	GluTyrrThrCysAanGluGlyPheLeuLeuCluGlyAlaArgSerArqValCysLeuAla	2875
Dd	1271	GAATACACTTTCAMATGAAGGGTTCTTGCTTGAGGGAGCCAGAGTGGGGTTGTCTTGCC	1330
Qy	2876	AasnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
Dd	1331	AATGGAAAGTTGGATGGAGCCACTCCCGACTGTGTGTCTGTCAGATGTGCCACCCGCCCA	1390
Qy	2896	GlnLeualaAanGlyValThrGluGlyLeuAspTyrrGlyPheMethLysGluValThrPhe	2915
Dd	1391	CAACTGGCCNAATGGGTCGACGGAAGGCTCGACTATGCTTCATGAAGGAAGTAACATTC	1450
Qy	2916	HisCysHisGluGlyTyrrIleLeuHisGlyAlaProLysLeuthrCysGlnSerAspGly	2935
Dd	1451	CACCTGTCCAGGGGCTACATCTTCGACGGTGCTCCAAAACCTCACCTGTCTAGTCAGATGGC	1510
Qy	2936	AasnTrpAspAlaGluIleProLeuCysLysProValAasnCysGlyProProGluAspLeu	2955
Dd	1511	AAC TGGAAGTGGAGATATCTCTCTCTGTAAACAGTCACTGTGGACCTCTCTGAAGATCTT	1570
Qy	2956	AlaHisGlyPheProAasnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrrGlnCys	2975
Dd	1571	GCCCAATGGTTTCCCTAATGGTTTTTCTTTATTCATGGGGSCCATATACAGTATCAGTGC	1630
Qy	2976	PheProGlyTyrrLysLeuHisGlyAasnSerSerArgArgCysLeuSerAasnGlySerTrp	2995
Dd	1631	TTTTCTGGTTATAAGCTCCATGGAAATTCATCAAGAAGGTGCCTCTCTCCAATGGCTCTCGG	1690
Qy	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrrGly	3015
Dd	1691	AGTGGCAGCTCACCTTCTGCTGCTCCCTTGCAGATGTTCCACACGAGTAATGAATATGGA	1750
Qy	3016	ThrValAanGlyThrAspPheAspCysGlyLysAlaalaArgIleGlnCysPheLysGly	3035
Dd	1751	ACTGTCNAATGGGACAGATTTTGA CTGTGGAAAGGCAGCCCGGATTCAGTGTCTTCAAAGGC	1810
Qy	3036	PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly	3055
Dd	1811	TTCAGCTCTTAGGACTTCTTCTGAAATCACCTGTGAAGCCGATGGAGCTGGAGCTCTGGG	1870

Qy	3056	PheProHisCysGluHisIsthSerCysGlySerLeuProMetIleProAsnAlaPheIle	3075
Db	1871	TTCCCCACCTGTGAACACACACTTCTTGTTGGTTCTCTTCCAATGATACCAATAGCGTTTCATC	1930
Qy	3076	SerGluThrSerSerTrpGlyAasnValIleThrTyTyrSerCysArgSerGlyTyVal	3095
Db	1931	AGTGAGACCCAGCTCTTTGGAAAGGAATATGATAACTTACAGCTGCAGGTTCTGGATATGTC	1990
Qy	3096	IleGlnGlySerSerAspLeuIleCysThrGluGlyValTrpSerGlnProTyrPro	3115
Db	1991	ATACAAGCCAGTTTCAGATCTGATTTTGTACAGAAAGGGGTATGGAGCCAGCCTTATCCA	2050
Qy	3116	ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly	3135
Db	2051	GTCTGTGAGCCCTTGTCCTGTGGGTCCCCACCGCTCTGTCCCAATGCAGTGGCACTGGA	2110
Qy	3136	GluAlaHisIsthTrpGluSerGluValValLeuArgCysLeuGluGlyTyThrMetAsp	3155
Db	2111	GAGGCACACACCTATGAAAGTGAAGTGAACACTCAGATGTCCTGGAAGGTTATACGATGGAT	2170
Qy	3156	ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSer	3175
Db	2171	ACAGATACAGATACATTCACCTCTGTAGAAAGATGGTGGTGTCTCCTGAGAGAACTCC	2230
Qy	3176	CysSerProLysCysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp	3195
Db	2231	TGCAGTCCCTAAAAATGTCCTCTCCCGGAAAAACAATACACATATACTTGTATCATGGGAC	2290
Qy	3196	AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyThrPheGluGly	3215
Db	2291	GATTTTCAGTGTGAATAGGCAGGTTTCTGTGTCAATGTCAGAAAGGGTATACCTTTGAGGGA	2350
Qy	3216	ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer	3235
Db	2351	GTTAACATATCAGTATGTAGCTGTATGTAAGTCTGGAGGCCACCATCTCCGATGATCT	2410
Qy	3236	CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer	3255
Db	2411	TGCAGTCCAGTTCTTGTGGGAAACCTCGAAAGTCCAGAACATGGAATTTGTGTTGGCAGT	2470
Qy	3256	LysTyThrPheGluSerThrIleIleTyGlnCysGluProGlyTyTrpGluLeuGluGly	3275
Db	2471	AAATACACCTTTGAAAGCACAAATATTATTATCACTGTGAGCCTGGCTATGAACTAGAGGGG	2530
Qy	3276	AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLys	3295
Db	2531	AACAGGGAACTGTCTGCCAGGAGACACAGACAGTGGAGTGGGGGTGGCAATATGCNAA	2590
Qy	3296	GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg	3315
Db	2591	GAGACCAGGTGTGAACACTCCACTTGAATTTCTCAATGGGAAGCTGACATTTGAAACACAG	2650
Qy	3316	ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyTrpSerLeuGluGlyProSer	3335
Db	2651	ACGACTGCAGCCCAACCGTGGTATATCTCTGCAACAGAGGCTTACAGTCTTGAAGGGCCATCT	2710
Qy	3336	GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn	3355
Db	2711	GAGGCACACTGCACAGAAATGGAACCTGAGCCACCCAGTCCCTCTCTGCAAAACCAAT	2770
Qy	3356	ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr	3375
Db	2771	CCATGCCCTGTCTCTTTTGTGATTCCCGAGATGCTCTGCTGCTGAAAGGAGGATTTTAT	2830
Qy	3376	ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle	3395
Db	2831	GTTGATCAGAAATGTGTCCATCAATATGTAGGGAAGGTTTCTGTCTGCAGGGCCACGGCATC	2890
Qy	3396	IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGlyLysIleSer	3415
Db	2891	ATTACCTGCACACCCCGACGACGTGCACACAGCAGCCCAATATGTGAAAATCTCA	2950
Qy	3416	CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyGlnTyGly	3435

Db 2951 TGTGGTCCACCAAGCTCAGGTAGAAAATGCAATGTCTCGAGGCGTACATATCAATATGGA 3010
Qy AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
Db 3011 GACATGATCACTACTCATGTACAGTGGATACATGTTGGAGGGTTTCTCGAGGAGTGT 3070
Qy CysLeuGluAenGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
Db 3071 TGTTTAGAAAATGGAAACATGAGACATCACCTCCTATTGTGACAGAGCTGTCTGTGCGATTTCCA 3130
Qy CysGlnAenGlyGlyIleCysGlnArgProAenAlaCysSerCysProGluGlyTrpMet 3495
Db 3131 TGTGAGATGGGGGCATCTGCCAAGCCCAATGCTGTCTCTCGAGGCGCTGGAGT 3190
Qy GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAenGlyGlyArgCysVal 3515
Db 3191 GGGCGCCTCTGGAAGAACCAATCTGCATTTCTCCCTGCTCGAAGGAGTGGTGTG 3250
Qy AlaProTyrGlnCysAaspCysProProGlyTyrTrpThrGlySerArgCysHisThrAlaVal 3535
Db 3251 GCCCTTACCACTGTGACTGCCCGCTGGCTGGACGGGTCTCGCTGTCTATACAGCTGT 3310
Qy CysGlnSerProCysLeuAenGlyGlyIleCysValArgProAenArgCysHisCysLeu 3555
Db 3311 TGCCAGTCTCCCTGCTTAATGGTGAATAATGTGAAGACCAACCGATGTCACTGTCT 3370
Qy SerSerTrpThrGlyHisAenCysSerArgGlySerArgTrpThrGlyPhe 3571
Db 3371 TCTTCTGGACGGACATAACTGTTCCAGAAAAGGAGGACTGGGT 3418

RESULT 18

US-09-764-881-21
; Sequence 21, Application US/09764881
; Publication No. US20020086821A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-881-21

Alignment Scores:
Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservatives: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 13 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-764-881-21 (1-3804)

Qy 2436 ValProValGluCysProGlnProGluGluIleProAenGlyIleIleAaspValGlnGly 2455
Db 11 GTTCAGTAGAATGTCCTGAGAAATCCCAATGGAATCATTTGATGTGCNAGGC 70
Qy 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAen 2475
Db 71 CTTCAGTATCTCAGCACAGCTCTATACCTGCAAGCCAGGCTTTGAAATGGTGGAAT 130
Qy 2476 ThrThrThrLeuCysGlyGluAenGlyHisTrpLeuGlyGlyIleProThrCysLysAla 2495
Db 131 ACTACACCCCTTGTGGAGAAAATGCTACTGCTTGGAGGAGAAACCAACATGTAAAGCC 190
Qy 2496 IleGluCysLeuLysProLysGluIleLeuAenGlyLysPheSerTyrThrAaspLeuHis 2515

Db 191 ATTGAGTGGCTGAACCCAGAGGATTTTGATGGCAATTTCTCTACAGGACCTACAC 250
Qy TyrGlyGlnThrValThrTyrSerCysAenArgGlyPheArgLeuGluGlyProSerAla 2535
Db 251 FATGACAGACGGTTACTTCTTGCAACAGAGGCTTTGGCTCGAAGGCTCCAGTGCC 310
Qy LeuThrCysLeuGluThrGlyAaspTrpAaspValAaspAlaProSerCysAenAlaIleHis 2555
Db 311 TTGACTGCTTTAGAGACAGGTGATTTGGGATGTAGATGGCCCATCTTTGCAATGCAATCCAC 370
Qy CysAaspSerProGlnProIleGluAenGlyPheValGluGlyAlaAaspTyrSerTyrGly 2575
Db 371 TGTGATTTCCCAACCAACCATTCGAAATGGTTTGTAGAGGTGCAGATTACAGCTATGGT 430
Qy AlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
Db 431 GCCATAATCATCTACAGTTGCTTCCCTGGGTTCAGGTGGCTGTCATGCGCATGCGAGCC 490
Qy CysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIleAaspCysGlyLeu 2615
Db 491 TGTGAAGAGTCAAGGATGGTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCCTC 550
Qy ProProHisIleAaspPheGlyAaspCysThrLysLeuLysAaspGlnGlyTyrPheGlu 2635
Db 551 CTTCTCTCATATGATTTTGGAGACTGTACTAACTCAAGATGACCCAGGGATATTTTGAG 610
Qy GlnGluAaspAaspMetMetGluValProTyrValThrProHisProProTyrHisLeuGly 2655
Db 611 CAAGAAGACGACATGATGGAAGTTCCATATGTGACTCTCCACCTCTCTATCAATTTGGGA 670
Qy AlaValAlaLysThrTrpGluAenThrLysGluSerProAlaThrHisSerSerAenPhe 2675
Db 671 GCAGTGGCTTAAACCTTGGGAAAATACAAAGGAGTCTCTGCTACACATTCATCAAACTTT 730
Qy LeuTyrGlyThrMetValSerTyrThrCysAenProGlyTyrGluLeuLeuGlyAenPro 2695
Db 731 CTGTATGTTACCATGTTTCATACACTGTATTCAGGATATGAACCTCTGGGGAACCT 790
Qy ValLeuIleCysGlnGluAaspGlyThrTrpAenGlySerAlaProSerCysIleSerIle 2715
Db 791 GTGCTGATCTGCCAGGAAGATGGAACCTTGGAAATGGCAGTGCACCATCTCGCATTTCAAT 850
Qy GluCysAaspLeuProThrAlaProGluAenGlyPheLeuArgPheThrGluThrSerMet 2735
Db 851 GAATGTGACTTGCCCTACTCTCTGAAATGGCTTTTTCGCTTTTACAGAGACTAGCATG 910
Qy GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAaspLeuArg 2755
Db 911 GGAAGTGTCTGTCAGTATAGCTGTAAACCTGGACACATCTAGAGGCTCTGACTTAAGG 970
Qy LeuCysLeuGluAenArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys 2775
Db 971 CTTTGTCTAGAGATAGAAAGTGGAGTGGTGGCTTCCCGCTGCTGGAAGCCATTTATGTC 1030
Qy LysLysProAenProValMetAenGlySerIleLysGlySerAenTyrThrTyrLeuSer 2795
Db 1031 AAAAGCCCAATCCAGTCTATGATGGAATCCATCAAGGAAGCACTACACATACCTGAGC 1090
Qy ThrLeuTyrTyrGluCysAaspProGlyTyrValLeuAenGlyThrGluArgArgThrCys 2815
Db 1091 ACGTTGTACTATGAGTGTGACCCCGGATATGTGCTGAATGGCACTGAGAGGAGACATGC 1150
Qy GlnAaspAaspLysAenTrpAaspGluAaspGluProIleCysIleProValAaspCysSer 2835
Db 1151 CAGGATGACAAAACCTGGGATGAGGATGAGCCATTTGCAATTCCTGTGGACTCGAGTTCA 1210
Qy ProProValSerAenGlyGlnValArgGlyAaspGlyThrPheGlnLysGluIle 2855
Db 1211 CCCCCAGTCTCAGCCATGCCCCAGGTGAGAGGAGCAGGTACACATTCAAAAAGAGATT 1270
Qy GluTyrThrCysAenGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla 2875

1271 GAATACACATTGCAATGAAGGTTCTTGCTTGAGGAGCCAGGAGTCGGGTTTGTCTTGCC 1330
2876 AsnGlySerThrProSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895
1331 AATGGAAGTTGGAGTGGAGCACTCCCGACTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1390
2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe 2915
1391 CAATGGCCATATGGGGTGAAGAGGCTGTGACTATGGCTTCATGAGAGGAGTAACTTC 1450
2916 HisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly 2935
1451 CACTGTCCAGAGGCTACATCTTTCGACGGTGTCCCAAACTCACTGTCTGTCTGTCTGTCT 1510
2936 AsnThrAspAlaGluIleProLeuLysLysProValAsnCysGlyProProGluAspLeu 2955
1511 AACTGGGATGCAGAGATTCCTCTCTGTAAACCAAGTCAACTGTGTGAGCTCTCTGAATCTT 1570
2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys 2975
1571 GCCCATGGTTTCCCTAATGGTTTTCTTTATTCATGGGGGCCATATACAGTATCAGTGC 1630
2976 PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerThr 2995
1631 TTTCTCTGGTTATAAGCTCCATGGAATTCATCAAGAAGGTGCCTCTCCAAATGCTCTCTGG 1690
2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
1691 AGTGGCAGCTCACT 1750
3016 ThrValAsnGlyThrAspPheAspCysGlyValAlaAlaArgIleGlnCysPheLysGly 3035
1751 ACTGTCAATGGGACAGATTTGACTGTGGAAGGACGCCCGAATTCAGTCTCTCAAGGC 1810
3036 PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnThrSerSerGly 3055
1811 TTCAAGCTCTTAGGACTTTCTGAATCACCTGTGAAGCGAGTGCCAGTGGAGCTCTGGG 1870
3056 PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
1871 TTCCCCCAGCTGTGAACACACTCTCTGTGGTTCCTTCTCAATGATACCAATGGGTTCATC 1930
3076 SerGluThrSerSerThrLysGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
1931 AGTGAGACCACTCTCTGGAGGAAATGTGATTAATCTTACAGCTGAGGCTGTGATATGTC 1990
3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValThrSerGlnProThrPro 3115
1991 ATACAAGGCAGTTTCAGATCTGATTTGTACAGAAAGGGGTATGGAGCCAGCCTTATCCA 2050
3116 ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly 3135
2051 GTCTGTGAGCCCTTGTCTGTGGGTCCCACTGTCTGTGGCAATGAGTGGCAACTGGA 2110
3136 GluAlaHisThrTyrGlnSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
2111 GAGGCAACACCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2170
3156 ThrAspThrAspThrPheThrCysGlnLysAspGlyArgThrPheProGluArgIleSer 3175
2171 ACAGATACAGATACATTCACCTGTGCAGAAAGATGGTGTCTGTCTGTCTGTCTGTCTCTCC 2230
3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
2231 TGCAGTCTCTAAAAATGTCTCTCCCGGAAACATAAACCATATATCTGTGTACATGGGAC 2290
3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
2291 GAATTCAGTGTGATAGGCAGATTTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2350
3216 ValAsnIleSerValCysGlnLeuAspGlyThrThrProProPheSerAspGluSer 3235
2351 GTTAAACATATCATGTATGTAGT 2410

3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValCysSer 3255
2411 TGCAGTCCAGTTCTTGTGTGGAAACCTGAAAGTCCAGAACTGGAATTTGTGTGTGGCAGT 2470
3256 LysTyrThrPheGluSerThrIleLeuTyrGlnCysGluProGlyTyrGluLeuGluGly 3275
2471 AATACACCTTTGAAAGCACAAATTTATTCAGTGTGAGCTGTGCTATGAAGAGGG 2530
3276 AsnArgGluArgValCysGlnGluAsnArgGlnThrSerGlyGlyValAlaIleCysLys 3295
2531 AACAGGGAAGTGTCTGCCAGAGAAACAGACAGTGGAGTGGAGGGTGGCAATATGCAAA 2590
3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAsnIleGluAsnArg 3315
2591 GAGACCAAGTGTGAACCTCCACTTGAATTTCTCAATGGGAAAGCTGACATTTGAAACAGG 2650
3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
2651 ACAGCTGGACCAACGTTGTATATTTCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCT 2710
3336 GluAlaHisCysThrGluAsnGlyThrThrSerHisProValProLeuCysLysProAsn 3355
2711 GAGGCACACTGCACAGAAATGGAACTGGAGGCCACCCAGTCCCTCTCTGCAAAACCAAT 2770
3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr 3375
2771 CCATGCCCTTCTCTTTGTGATTCGCGAATGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 2830
3376 ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
2831 GTTGATCAGAAATGTGTCCATCAATGTAGGAAAGTTTCTGTCTGAGGGCCACGGCATC 2890
3396 IleThrCysAsnProAspGluThrThrThrGlnThrSerAlaLysCysGluLysIleSer 3415
2891 ATTACCTGCAACCCCGACGACGTGGACACAGACAGCAAGCGCCAAATGTGAAAAATCTCA 2950
3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
2951 TGTGTCTCACCACTCACGTAGAAATATGCAATGTCTGAGCGGTACATTTATCAATATGGA 3010
3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
3011 GACATGATCACCTACTCATGTACAGTGGATACATGTTGGAGGGTTTCTCTGAGGAGTGT 3070
3456 CysLeuGluAsnGlyThrThrThrSerProProIleCysArgAlaValCysArgPhePro 3475
3071 TGTTTAGAAAAATGGAACATGGACATCACCTCTCTATTTCAGAGCTGTCTGTCTGATTTCCA 3130
3476 CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMet 3495
3131 TGTGAGAAATGGGGGATCTGCCAACGCCCAATATGCTTGTCTGTCTGAGAGGGCTGGATG 3190
3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
3191 GGGGGCTCTGTGAAGAACCAATCTGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3250
3516 AlaProTyrGlnCysAspCysProProGlyTyrThrThrGlySerArgCysHisThrAlaVal 3535
3251 GCCTCTTACAGTGTGACTGCCCGCTGTGTCGAGCGGGTCTGCTGTCTGTCTGTCTGTCTGT 3310
3536 CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu 3555
3311 TGCAGTCT 3370
3556 SerSerThrThrGlyHisAsnCysSerArgLysArgThrGlyPhe 3571
3371 TCTTCTTGGACGGACATAACTGTTTCAGGAAAGGAGGACTGGTTT 3418

RESULT 19

US-09-764-893-43

; Sequence 43, Application US/09764893

; Publication No. US20020086330A1

; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: P0209
 ; CURRENT APPLICATION NUMBER: US/09/764,893
 ; CURRENT FILING DATE: 2001-01-19
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 154
 ; SOFTWARE: PatentIn ver. 2.0
 ; SEQ ID NO 43
 ; LENGTH: 3804
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-893-43

Alignment Scores:
 Pred. No.: 0 Length: 3804
 Score: 6452.00 Matches: 1135
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.91% Mismatches: 0
 Query Match: 32.30% Indels: 0
 DB: 13 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-764-893-43 (1-3804)

QY	2436	ValProValGluCysProGlnProGluGluLeuProAsnGlyIleLeuAspValGlnGly	2455
DB	11	GTTCAGTAGAATGTCCTCCCAACCTGAGAAATCCCAATGGAATCATTTGATGTGCAAGGC	70
QY	2456	LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsn	2475
DB	71	CTTGCTTACTCTGACGACAGCTCTCTATACCTGTCAGCCAGGCTTTGANTTGGTGGGAAT	130
QY	2476	ThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAla	2495
DB	131	ACTACCACCTTTGTGGAGAAATGGTCTACTGGCTTGAGGAAACCAACATGTAAAGCC	190
QY	2496	IleGluCysLeuLysProLysGluLeuAsnGlyLysPheSerTyrThrAspLeuHis	2515
DB	191	ATTGAGTGCTCCGAAACCCCAAGAGATTTTGAATGGCAATTTCTCTTACAGGACCTACAC	250
QY	2516	TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla	2535
DB	251	TATGGACAGACCGTTTACTACTCTTCTGCAACCGAGGCTTCGGCTCGAAGGTCCCAAGTCC	310
QY	2536	LeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHis	2555
DB	311	TTGACCTGTTTAGACACAGGTGATTGGGATGTAGATGCCCATCTTGCATATGCCATCCAC	370
QY	2556	CysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGly	2575
DB	371	TGTGANTCCCAACACCAATTTGAAATGGTTTGTAGAAGGTGACAGATTTACAGCTTATGGT	430
QY	2576	AlaIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr	2595
DB	431	GCCATAATCACTACAGTTGCTTCTCTGGTTTCAGGTGGCTGGTTCATGCCATGCAGACC	490
QY	2596	CysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeu	2615
DB	491	TGTGAAGAGTCAGGATGGTCAAGTTCATCCCAACATGTATGCCAATAGACTGTGGCCCTC	550
QY	2616	ProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGlu	2635
DB	551	CCTCTCATATAGATTTTGGAGACGTGTACTAACTCAAGATGACAGGATATTTTGGAG	610
QY	2636	GlnGluAspAspMetGluValProTyrValThrProHisProProTyrHisLeuGly	2655
DB	611	CAAGAGACGACATGATGGAAGTTCATATATGACTCTCTCCCTCTTATCATTTTGGGA	670
QY	2656	AlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhe	2675
DB	671	GCAGTGGCTAAACCTGGGAAATCAAAAGGAGTCTCTGTGTACACATTTTCAATCAACTTT	730

QY	2676	LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnPro	2695
DB	731	CTGTATGGTACCATGGTTTTCATACCTGTATTCAGGATATGAACCTTCTGGGAAACCT	790
QY	2696	ValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle	2715
DB	791	GTGCTGATCTGCCAGGAAGATGGAACTTGGAAATGGCAGTGCCACCATCTCTGCAATTT	850
QY	2716	GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMet	2735
DB	851	GAATGTGACTTGTGCTCTCTGCTGCTCTCTGCTGCTCTCTGCTGCTCTCTGCTGCT	910
QY	2736	GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg	2755
DB	911	GGAGTGTCTGTCAGTATAGCTGTAACTGGACACATCTTAGCAGGCTCTGACTTAAGG	970
QY	2756	LeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys	2775
DB	971	CTTTGTCTAGAGATAGAAAGTGGAGTGGTCTCTCCCAACGCTGTGAAGCCATTTTCATGC	1030
QY	2776	LysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSer	2795
DB	1031	AAAAAGCCAAATCCAGTTCATGATGATCCATCAAGGAAGCACTACACATACCTGAGC	1090
QY	2796	ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCys	2815
DB	1091	ACGTTGTTACTAGTGTGACCCCGGATATGTCTGATGGCACTGAGAGGAGACATGC	1150
QY	2816	GlnAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer	2835
DB	1151	CAGGATGACAAACCTGGGATGAGGATGAGCCATTTGCAATTTCTGTGACTGCAATTC	1210
QY	2836	ProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle	2855
DB	1211	CCCCAGGTCTCAGCCAAATGGCCAGGTGAGAGACGAGTACACATTCCTCAAAAGAGATT	1270
QY	2856	GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla	2875
DB	1271	GAATACACTTGCATGATGAGGTTCTTGTCTGAGGAGCCAGGAGTCGGTGTCTTGTCC	1330
QY	2876	AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
DB	1331	AATGGAGTTGGAGTGGAGCCACTCCGACTGTGTGCTGTGATGATGTGCCACCCGCCA	1390
QY	2896	GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe	2915
DB	1391	CAACTGGCCAAATGGGGTCAACGAGGCTTGGACTATGGCTTCATGAGGAGTAACTTC	1450
QY	2916	HisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly	2935
DB	1451	CACTGTCAAGAGGGCTACATCTTGACCGTGTCTCCAAACCTCACCTGTCTGATCAGTGC	1510
QY	2936	AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProGluAspLeu	2955
DB	1511	AATGGGATGACAGATATCTCTCTGTAAACAGTCAACTGTGGACCTCTCAAGATCTT	1570
QY	2956	AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys	2975
DB	1571	GCCCATGGTTTCTTATGGTTTCTTATTCATGGGGCCATATACAGTATCAGTGC	1630
QY	2976	PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp	2995
DB	1631	TTTCTGTGTATAGCTCCATGGAATTCATCAAGAGGTGCTCTCTCAATGGCTCTCTGG	1690
QY	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly	3015
DB	1691	AGTGGCAGCTCACCTTCTCTGCTGCTTGCAGATGTTCACACCACTAGTAATTAATGGA	1750
QY	3016	ThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly	3035
DB	1751	ACTGTCAATGGGACAGATTTTACCTGTGGAAGGACGCCCGGATTCAGTGTCTCAAGGC	1810
QY	3036	PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly	3055

1811 TTCAAGCTCTAGGACTTTCTGAAATCACCTGTGTAGAGCCGATGCGCAGTGTGAGCTCTGGG 1870
3056 PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
1871 TTCCCCACCTGTGAACACACTTCTGTGGTCTCTTCCAAATGATACCAATGGTTCATC 1930
3076 SerGluThrSerSerTrpIysGluAsnValIleThrTyrSerCysArgSerGlyVal 3095
1931 AGTGAGACCCAGCTTCTTGGAGGAAATGTGTAACCTTACAGCTGCAGGTCTGGATATGTC 1990
3096 IleGlnGlySerSerAspLeuIleCysThrGluYsGlyValTrpSerGlnProTyrPro 3115
1991 ATACAAGGCAGTTCCAGATCTGAATTTGTACAGAGAGAGGGGTATGGACCGCTTATCCA 2050
3116 ValCysGluProLeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGly 3135
2051 GTCTGTGAGCCCTGTCTCTGGGTCCCAACCTCTGTGCGCAATGCAATGGCACTGGA 2110
3136 GluAlaHisThrTyrGluSerGluValIleValLeuArgCysLeuGluGlyTyrThrMetAsp 3155
2111 GAGGCAACACACTTATGAAGTGAAGTGAACCTCAGATGTCTGGAAGTTTATACGATGGAT 2170
3156 ThrAspThrAspThrPheThrCysGlnYsAspGlyArgTrpPheProGluArgIleSer 3175
2171 ACAGATACAGATACATTCACCTGTCCAGAAAGATGGTCTGCTGCTCCCTGAGAGATCTCC 2230
3176 CysSerProIysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
2231 TGCAGTCTTAAAAATGTCTCTCCCGAAAAACATACACATATCTTGTACATGGGGAC 2290
3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
2291 GAATTTCAAGTGAATAGGCAAGTTCTGTGTCTATGTGCAAGGGTATACCTTTGAGGGA 2350
3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSer 3235
2351 GTTAACATATCAGTATGTCAGCTTGTATGGAACCTGGGAGCCACCATCTCTCCGATGATCT 2410
3236 CysSerProValSerCysGlyYsProGluSerProGluHisGlyPheValValGlySer 3255
2411 TGCAGTCCAGTCTTGTGTGGAAACCTGAAAGTCCAGAACATGGATTTGTGGTGGCAGT 2470
3256 IysThrThrPheGluSerThrIleTyrGlnCysGluProGlyTyrGluLeuGluGly 3275
2471 AAATACACCTTTGAAGACCAATTTATTTATCAGTGTGAGCTGGCTGATGAACATAGAGGG 2530
3276 AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLys 3295
2531 AACAGGGAACGTGTCTCCAGGAGAGACAGACAGTGGAGGTGGCAATATGCATA 2590
3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg 3315
2591 GAGACCGGTGTGAACCTCCACTTGAATTTCTCAATGGGAAAGCTGACATTTGAACACAGG 2650
3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
2651 ACAGCTGGACCCCAACGTGTATATCTCTGCAACAGAGGCTTACAGTCTTGAAGGGCCATCT 2710
3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn 3355
2711 GAGGCACTGTCACAGAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGCACCAACCAAT 2770
3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluYsGluPheTyr 3375
2771 CCATGCCCTGTCTCTTTGTGATTTCCCGAGAAATGCTCTGTCTGTCTCAAAAGGAGTTTAT 2830
3376 ValAspGlnAsnValSerIleYsCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
2831 GTTGATCAGAAATGTGTCCATCAATGTAGGAGGTTTCTGTCTGAGGCGCCACCGCATC 2890
3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluValIleSer 3415

2891 ATTACCTGCAACCCCGACGAGACGCTGACACACAGACGCGCAAAATGTGAAAAATCTCA 2950
3416 CysGlyProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
2951 TGTGGTCCACAGCTCACGTAGAAAAATGCAATTTCTCGAGGCGTACATTATCAATATGGA 3010
3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
3011 GACATGATACCTACTCATGTTCAGTGGATACATGTTGGAGGGTTTCTCTGAGGAGTGT 3070
3456 CysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
3071 TGTTTAGAAAAATGGAACATGACATCACCTCTATTTTCAGAGCTGTCTGTGCAATTTCCA 3130
3476 CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMet 3495
3131 TGTGAGAAATGGGGCATCTGCCAACGCCCAATGCTTCTCTGTCAGAGGGCTGGAGT 3190
3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
3191 GGGCGCTCTGTGAGAGAACCAATCTGCATTTCTCTGTCGACGGAGGTGCTGTGTG 3250
3516 AlaProTyrGlnCysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
3251 GCCCTTACAGTGTGACTGCCGCTGGCTGGAGCGGGTCTCGCTGTCTACAGCTGT 3310
3536 CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu 3555
3311 TGCAGCTCTCCCTGCTTAAATGGTGGAAATGTGTAAAGCCAAACCGATGTCTGTCT 3370
3556 SerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe 3571
3371 TCTCTTGGACGGGACATAACTGTTCAGGAAAGAGGAGGACTGGGTTT 3418

RESULT 20
US-10-073-865-43
; Sequence 43, Application US/10073865
; Publication No. US20030044904A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: FUZ09C1
; CURRENT APPLICATION NUMBER: US/10/073,865
; CURRENT FILING DATE: 2002-02-14
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-073-865-43

Alignment Scores:
Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 15 Gaps: 0

US-09-977-053-4 (1-3571) x US-10-073-865-43 (1-3804)

2436 ValProValGluCysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGly 2455
11 GTTCCAGTAGAATGTCCCAACCTCGAAGAAATCCCAATGGAATCATTTGATGTGCAAGC 70
2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsn 2475
71 CTGGCTATCTCAGCACAGCTCTCTATCTGACCTGCAAGCAGGCTTGAATGTGGGAAT 130
2476 ThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAla 2495

131 ACTACCAACCTTTGTGGAGAAAATGCTCACTGGCTTGGAGGAAAAACCAACATGTAAAGCC 190
2496 IleGluCysLeuIysProLysGluIleLeuAsnGlyLysPheSerThrAspLeuHis 2515
191 ATTGAGTGCCTGAACCCCAAGGAGATTTTGAATGGCAAAATTCCTTTACACGGACCTTACAC 250
2516 TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla 2535
251 TATGGACAGACGTTTACCTACTCTTCCAAACCGAGGCTTTCGGCTCGAAGGTCCAGTGGCC 310
2536 LeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHis 2555
311 TTGACCTGTTTAGACAGACGTCATTTGGATGTAGATGCCCACTCTTGTCAATGCCATCCAC 370
2556 CysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerThrGly 2575
371 TGTGATTCCTCCCAACCCCAATGAAATGGTTTGTAGAAGGTGCAGATTACAGCTATAGTT 430
2576 AlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
431 GCATATATCATCTACAGTTGCTTCCCTGGGTTTCAGTGGCTGCTCATGCCATGCCAGACC 490
2596 CysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeu 2615
491 TGTGAAGAGTCAGGATGTCCTCACTCCCAACATGTATGCCAATAGACTGTGGCCCTC 550
2616 ProProHisIleAspPheGlyAspCysThrLysLeuIysAspAspGlnGlyTyrPheGlu 2635
551 CCTCCTCATATAGATTTTGGAGACTGTACTAACTCAAGATGACACAGGGATATTTTGA 610
2636 GlnGluAspAspMetGluValProTyrValThrProHisProTyrHisLeuGly 2655
611 CAAGAAGACACATGATGAGTTCCATATGTGACTCTCCACCTCTTATCATTTGGGA 670
2656 AlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhe 2675
671 GCAGTGGCTAAACCTGGGAAAATACAAGGAGTCTCTCTGTACACATTCATCAAACTTT 730
2676 LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnPro 2695
731 CTGTATGGTACCATGGTTTTCATACACTGTAAATCCAGGATATGAATCTTCTGGGAAACCT 790
2696 ValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle 2715
791 GGTCTGATCTGCCAGGAAGATGGAATCTTGGATGGCAGTGCACCATCTCTGCATTTCAAT 850
2716 GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMet 2735
851 GAATGTGACTTGCTTACTGCTCTCGAAAATGGCTTTTGGCTTTTACAGAGACTAGCATG 910
2736 GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg 2755
911 GGAAGTGTGTGCAGTATAGCTGTAAACCTGGACACATTTCTAGACGCTCTGACTTAAGG 970
2756 LeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys 2775
971 CTTTGTCTAGAGATAGAAAGTGGAGTGGTGGCTCCCAAGCTGTGGAGCCATTTCAATGC 1030
2776 LysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSer 2795
1031 AAAAGCCCAATCCAGTCAATGAATGATCCATCAAGAGGAAGCAATACACATACCTGAGC 1090
2796 ThrLeuTyrTyrGlnCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCys 2815
1091 ACGTTGTACTATGAGTGTGACCCCGATATGTGCTGAATGGCACTGAGAGGAGAACATGC 1150
2816 GlnAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer 2835
1151 CAGGATGACAAAACTGGGATGAGGATGAGCCCATTTGGCATTTCTGTGGACTGCACTTCA 1210
2836 ProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle 2855
1211 CCCCCAGTCTCAGCCAAATGCCAGGTGAGAGGAGACGAGTACACATTCAAAAAGAGATT 1270

2856 GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyValaArgSerArgValCysLeuAla 2875
1271 GAATACACTTGCATAGAGGTTCTTGTCTGAGGAGCCAGGAGTCGGGTGTGCTTGTGCC 1330
2876 AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895
1331 AATGGAAGTTGGATGGAGCCACTCCGACTGTGTGCTGTCTCAGATGTGTCACCCGCCA 1390
2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe 2915
1391 CAACCTGCCCAATGGGGTGACGGAAGCCTCGCATATGGCTTTCATGAAGGAAGTAACATTC 1450
2916 HisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly 2935
1451 CACTGTACAGAGGCTACATCTTGCACGGTCTCCAAACTCACCTGTCTCAGTCAAGATGTC 1510
2936 AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeu 2955
1511 AACTGGATGTCAGAGATTCCTCTCTGTAAACCACTCACTGTGGACCTCTCTGAAGATCTT 1570
2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCys 2975
1571 GCCCATGGTTTCCCTAATGGTTTTTCTTTATTTATGAGGGGCCATATATACAGTATCAGTGC 1630
2976 PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp 2995
1631 TTTCTCTGGTTATAGCTTCCATGGAAATTCATCAAGAAGGTGCTCTCTCAATGGCTCTGG 1690
2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
1691 AGTGGCAGCTCACCTTCTCTGCTGCTTGCAGATGTTCCACACCACTGAATTAATATGGA 1750
3016 ThrValAsnGlyThrAspPheAspCysGlyValAlaAlaArgIleGlnCysPheLysGly 3035
1751 ACTGTCAATGGGACAGATTTTGCATGTGGAAGGACGACCCGGAATTCAGTGTCTCAAAAGC 1810
3036 PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly 3055
1811 TTTCAAGCTCTAGAGCTTCTGAAATCCTCTGTGAGCCGATGGCCAGTGGAGCTCTGGG 1870
3056 PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
1871 TTTCCCCACTGTGACACACACTTCTTGTGGTCTCTTCCAATGATACCAATGCGCTCATC 1930
3076 SerGluThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
1931 AGTGAGACCAAGCTCTTGGAGGAAAATGTGATAACTTTACAGCTGCAAGTCTGGATATGTC 1990
3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrPro 3115
1991 ATACAGGCAGTTTCAGATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCA 2050
3116 ValCysGluProLeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGly 3135
2051 GTCTGTGAGCCCTGTCTGTGGGTCCCAACCGCTCTGTGCCAATGCAATGCAAGTCACTGA 2110
3136 GluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
2111 GAGGCACACACTTATGAAGTGAAGTGAAGTCAAGATGTCGGAAGGTATACCATGGAT 2170
3156 ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSer 3175
2171 ACAGATACAGATACATTCACCTGTCTAGAAAGATGGTGGCTGGTTCCCTGTGAGAGATCTCC 2230
3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
2231 TGCAGTCTTAAAAAATGCTCTCTCCCGGAAAAACATAACACATATATCTGTATACATGGGAC 2290
3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
2291 GATTTTCAGTGTGAATAGGCAAGTTTCTGTGTCAATGTGAGAGGGGTATACCTTTTGAGGGA 2350

Qy	3216	ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer	3235
Db	2351	GTTAACACATATCAGTATGTCAGCTTGATGGAACCTGGGAGGCACCATTTCTCCGATGAATCT	2410
Qy	3236	CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer	3255
Db	2411	TGCAGTCCAGTTCTGTGGGAACCTGAAGTCCAGAACATGGATTTGTGTTGGCAGT	2470
Qy	3256	LysThrPheGluSerThrIleIlePyrGlnCysGluProGlyTyrGluLeuGluGly	3275
Db	2471	AAATACACCTTTTGAAGACCAATATTATTATCAGTGTGAGCTGGCTATGAATAGAGGG	2530
Qy	3276	AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLys	3295
Db	2531	AACAGGGAACGTGTCTCCACAGAGAAACAGACAGTGGAGTGGGGGGTGGCAATATGCAAA	2590
Qy	3296	GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg	3315
Db	2591	GAGACCAAGGTGTGAACCTCCACCTTGAATTTCTCAATGGGAAGAGCTGACATTTGAAACAGG	2650
Qy	3316	ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer	3335
Db	2651	ACGACTGGACCCAAACGTGGTATATTCCTGCACACAGAGGCTACAGCTTTGAAGGGCCATCT	2710
Qy	3336	GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn	3355
Db	2711	GAGGCACACTGCACAGAAATAGAACTGGAGCCACCAGTCCCTCTCTGCAAAACCAAT	2770
Qy	3356	ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr	3375
Db	2771	CCATGCCCTGTCTCTTTGTGATTTCCGAGAAATGCTCTGCTGTGAAAGGAGTATTTAT	2830
Qy	3376	ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle	3395
Db	2831	GTTGATCAGAAATGTGTCATCAAAATGAGGAGGTTTCTGCTGCAGGGCCACGGCATC	2890
Qy	3396	IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSer	3415
Db	2891	ATTACCTGCAACCCCGACGAGCGTGACACACAGCAAGCGCCAATGTGAAAAAATCTCA	2950
Qy	3416	CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly	3435
Db	2951	TGTGGTCCACACAGCTCAGTAGAAAAATGCAATTGCTCGAGGCGTACATATATATATGGA	3010
Qy	3436	AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal	3455
Db	3011	GACATGATCACCTACTCATGTTACGTGGATACATGTTGGAGGGTTTCTTGAGAGAGTGT	3070
Qy	3456	CysLeuGluAsnGlyThrTrpThrSerProPheCysArgAlaValCysArgPhePro	3475
Db	3071	TGTTTAGAAAAATGAAACATGACATCACCTCTATTTTGCAGAGCTGTCTGTGCAATTTCCA	3130
Qy	3476	CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMet	3495
Db	3131	TGTCAAGATGGGGGCATCTGCACAGCCCAAAATGCTTGTTCCTGTCCAGAGGGCTGGATG	3190
Qy	3496	GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal	3515
Db	3191	GGGCGCTCTGTGAAGAACCAATCTGCATCTTCTCCCTGTCTGAACGGAGTGGCTGTGTG	3250
Qy	3516	AlaProTyrGlnCysAspCysProProGlyTrpThrGlySerArgCysHisThrAlaVal	3535
Db	3251	GCCCTTTACCAATGTGATCGCCGCCCTGGCTGGACGGGGTCTGCTGTGTCTATACAGCTGTT	3310
Qy	3536	CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu	3555
Db	3311	TGCAGTCTCCCTGCTTAATGGTGGAAATATGTGTAGACCAACCAACCGATGTCACTGTCTT	3370
Qy	3556	SerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe	3571
Db	3371	TCCTTTTGGACGGACATAACTGTTCTCAGGAAAAAGGAGGACTGGGTTT	3418

```

US-10-242-747-21
; Sequence 21, Application US/10242747
; Publication No. US20040005577A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OP INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PIZ07C1
; CURRENT APPLICATION NUMBER: US/10/242,747
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,881
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 3804
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-242-747-21

Alignment Scores:
Pred. No.: 0 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 16 Gaps: 0

US-09-977-053-4 (1-3571) x US-10-242-747-21 (1-3804)
Qy 2436 ValProValGluCysProGlnProGluGluileProAsnGlyIleIleAspValGln
Db 11 GTTCCAGTAGAATGTGCCCCAACCTGAGAAATCCCCCAATGGAATCATTTGATGTGCAAT
Qy 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLeysProGlyPheGluLeuValGly
Db 71 CTTGSCCTATTCTCAGCACAGCTCTCTATACCTGCAAGCCAGGCTTTTGAATTTGGTGGGA
Qy 2476 ThrThrThrLeuCysGlyGluAsnGlyHisIlePLeuGlyGlyIlysProThrCysIlys
Db 131 ACTACCAACCCCTTTGTGGAGAAATGTCATCTGGCTTGGAGGAAACCAACATGTAA
Qy 2496 IleGluCysLeuIysProTysGluIleLeuAsnGlyIlysPheSerTyrThrAspLeu
Db 191 ATTGAGTGGCTGTAACCCCAAGGAGATTTTGAATGGCAAAATTCCTTTACACGACCTTA
Qy 2516 TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSer
Db 251 TATGGACAGACCGTGTACCTACTCTTTCGCAACCCGAGGCTTTTGGGCTCGAAGTCCCACT
Qy 2536 LeuThrCysLeuGluThrGlyAspTProAspValAspAlaProSerCysAsnAlaIle
Db 311 TTGACCTGTTTAGACACAGGTGATTTGGAGTGTAGATGCCCATCTTTGCAATGCAATGCATC
Qy 2556 CysAspSerProGlnProIleGluAsnGlyPheValGluIleAlaAspTyrSerTyr

```

Db 371 TGTGATTTCCCAACCCATTGAAATGTTTTGTAGAGGTGCAGATTACAGCTATGGT 430
Qy 2576 AlalilelleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr 2595
Db 431 GCCATAATCATCTACAGATTGCTTCCCTGGGTTTCAGGTGGCTGGTTCATGCCATGCAGACC 490
Qy 2596 CysGluGluSerGlyTyrSerSerSerIleProThrCysMetProIleAspCysGlyLeu 2615
Db 491 TGTGAAGAGTCAGAGTGGTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCCCTC 550
Qy 2616 ProProHisIleAspPheGlyAspCysThrLysLeuLysAspGlnGlyTyrPheGlu 2635
Db 551 CCTCTCATATAGATTTTGGAGACTGTACTAACTCAAGATGACCCGGAATATTTTGAG 610
Qy 2636 GlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHisLeuGly 2655
Db 611 CAAGAAGACGACATGATGGAAGTTCATATGTGACTCTCCCTCACCATTCATCAACATTT 670
Qy 2656 AlalAlaIleTyrThrGluSerThrLysGluSerProAlaThrHisSerSerAsnPhe 2675
Db 671 GCAGTGGCTAAACCTGGGAAATACAAAGGAGTCTCCCTGCTACACATTCATCAACATTT 730
Qy 2676 LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnPro 2695
Db 731 CTGTATGTACCATGGTTTCATACACCTGTATCCAGGATATGAACCTCTGGGGAACTCT 790
Qy 2696 ValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle 2715
Db 791 GTGCTGATCTGCAGAGATGGAATCTGGAATGGCAGTGCCACCATCTCTGCAATTTCAATT 850
Qy 2716 GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGlnThrSerMet 2735
Db 851 GAATGTGACTTGCCCTACTGCTCTCGAAATGGCTTTTTCGGTTTACAGAGACTAGCATG 910
Qy 2736 GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg 2755
Db 911 GGAAGTGTGTGCACTATAGCTGTAACTTGACACATTCAGCAGGCTCTGACTTAGG 970
Qy 2756 LeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCys 2775
Db 971 CTTTGTCTAGAGATAGAAAGTGGAGTGGTGGCTCCCAACGCTGTGAAGCCATTTTCATGC 1030
Qy 2776 LysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSer 2795
Db 1031 AAAAGGCCAAATCCAGTCAATGAATGATCCATCAAGGAAGCAACTACACATACCTGAGC 1090
Qy 2796 ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCys 2815
Db 1091 AGTTGTACTATGAGTGTGACCCCGGATATGTGCTGAATGGCACTGAGAGGAGACATGC 1150
Qy 2816 GlnAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer 2835
Db 1151 CAGGATGACAAAACTGGGATGGAGTGAAGCCCATTTGCAATTCCTGTGGACTGCAGTTCA 1210
Qy 2836 ProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIle 2855
Db 1211 CCCCCAGTCTCAGCCAAATGGCCCGGTGAGGAGAGCAGTACACATTCNAAAAGAGATT 1270
Qy 2856 GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla 2875
Db 1271 GAATACACTTGCATGAGGGTTCTTGCTTGAGGGAGCCAGGAGTCGGTGTGCTTGCC 1330
Qy 2876 AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895
Db 1331 AATGGAAGTGGAGTGGAGCCACTCCGCACTGTGTGCTGTGCAGATGTGCCACCCGCCA 1390
Qy 2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPhe 2915
Db 1391 CAACCTGGCCAAATGGGGTGAACGAAGCCCTGGACTATGGCTTCATGAAGGAATACATTC 1450
Qy 2916 HisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly 2935

Db 1451 CACTGTGCAGAGGGCTACATCTTGCACGGTGTCTCCAAAACTCACCTGTCTAGTCAGATGGC 1510
Qy 2936 AsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeu 2955
Db 1511 AACTGGGATGCAGAGATTCCTCTCTGTAAACCAAGTCAACTGTGGACCTCTCTGAAGATCTT 1570
Qy 2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys 2975
Db 1571 GCCCATGGTTTCCCTTAATGGTTTTCTTTATTCATGGGGCCATATACAGTATCAGTGC 1630
Qy 2976 PheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp 2995
Db 1631 TTTCTCGTTATAAGCTCCATGCAATTCATCAAGAAGGTGCTCTCTCCAATGGCTCTCGG 1690
Qy 2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProAlleGluTyrGly 3015
Db 1691 AGTGGCAGCTCCACTTCTGCTGCTTGCAGATGTTCACACAGTAATTAATATGGA 1750
Qy 3016 ThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly 3035
Db 1751 ACTGTCAATGGGACAGATTTTGACTGTGGAAAGGACGCCCGGATTCAGTGTCTCAAGGC 1810
Qy 3036 PheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly 3055
Db 1811 TTCAAGCTCTTAGCACTTTCTGAAATCACTGTGAAGCCGATGGCCAGTGGAGCTCTGGG 1870
Qy 3056 PheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
Db 1871 TTCCCCCACTGTGAAACACACTTTCTGTGGTGTCTTTCCAAATGATACCAATGGCTTCATC 1930
Qy 3076 SerGluThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
Db 1931 AGTGAACACAGCTCTTGGAAAGGAAATGTGATAACTTACAGCTGCAGGTCTGGATATGTC 1990
Qy 3096 IleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrPro 3115
Db 1991 ATACAAGGCAGTTTCAGATCTGATTTGTACAGAAAGGGGTATGGAGCCAGCCCTTATCCA 2050
Qy 3116 ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly 3135
Db 2051 GTCTGTGAGCCCTGTGTCTGGGTCCCCCAGCTGTGTGCCAATGAGTGGACACTGGA 2110
Qy 3136 GluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAsp 3155
Db 2111 GAGGCACACACTATGAAGTGAAGTGAAGTCAAGATGTCAGATGTCGGAAGTTTATACATGGAT 2170
Qy 3156 ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSer 3175
Db 2171 ACAGATACAGATACATTCACCTGTCAAGAAAGATGGTGGCTGCTGCTCCCTGAGAGAATCTCC 2230
Qy 3176 CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
Db 2231 TGCAGTCTTAAAAAATGTCTCTCCCGAAACATTAACATATATCTTGTACATGGGAC 2290
Qy 3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
Db 2291 GATTTCAAGTGTGAATAGCAAGTTTCTGTGTCTCATGTGTCAGAGAGGTATACCTTTAGGGA 2350
Qy 3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer 3235
Db 2351 GTTAAACATATCAGTATGTCAGCTTGAATGAACCTGGGAGCCACCATTTCTCCGATGAATCT 2410
Qy 3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer 3255
Db 2411 TGCAGTCCAGTTTCTTGTGGAAACCTGAAGTCCAGAACATGGAATTTGTGGTGGCAGT 2470
Qy 3256 LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGly 3275
Db 2471 AAATACACTTTGAAGCACCAATTTATTTATCAGTGTGAGCCTGGCTATGAACCTAGAGGGG 2530
Qy 3276 AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLys 3295
Db 2531 AACAGGAAACGTGTCTGCCAGGAGAACACAGAGTGGAGTGGAGGGGTGGCAATATGCAAA 2590

QY 3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyValAlaAspIleGluAsnArg 3315
 DB 2591 GAGACCGAGTGTGAACCTCCACTTGAATTTCTCAATGGGAAGCTGACATTTGAACACAGG 2650
 QY 3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
 DB 2651 AGACTGGACCCACACGCTGTATATCTCTGCAACGAGGCTACAGTCTTGAGGGCCATCT 2710
 QY 3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuGlyCysValProAsn 3355
 DB 2711 GAGGCACACTGCACAGAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGCAACCAAT 2770
 QY 3356 ProCysProValProPheValIleProGluAsnAlaLeuSerGluLysGluPheTyr 3375
 DB 2771 CCATGCCCTTCTCTTTTGTGATTCCTGAGATGCTCTGCTGTCTGAAAGAGGATTTAT 2830
 QY 3376 ValAspGluAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
 DB 2831 GTTGATCAGATGTGTCATCAATGTAGGAAGGTTTCTGCTGAGGGCCACGGCCTC 2890
 QY 3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSer 3415
 DB 2891 ATTACTGCAACCCCGACGAGCTGCACACAGCAAGCGCCAAATGTGAATAATCTCA 2950
 QY 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
 DB 2951 TGTGCTCCACCACTCACGTAGAAAATGCAATTTGCTCGAGGCGTACATTAATATGGA 3010
 QY 3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
 DB 3011 GACATGATCACCCTACTCTTACAGTGGATACATGTTGAGAGGTTTCTCTGAGAGTGT 3070
 QY 3456 CysLeuGluAsnGlyThrTrpThrSerProPheIleCysArgAlaValCysArgPhePro 3475
 DB 3071 TGTAGTAAATGGAACATGACATGACATCCTCTATTTCAGAGCTGTCTGTGATTTCCA 3130
 QY 3476 CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMet 3495
 DB 3131 TGTGAGATGGGGGATCTGCCAACCCCAATGCTTTGCTGTCCAGAGGCTGTGATG 3190
 QY 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
 DB 3191 GGGCGCTCTGTGAAGAACCAATCTCATCTCTCTGCTGCTGAACGAGGCTGCTGTG 3250
 QY 3516 AlaProTyrGlnCysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
 DB 3251 GCCCTTACAGTGTGACTGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3310
 QY 3536 CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu 3555
 DB 3311 TGCCAGTCTCCCTGCTTAAATGGTGGAAATGTGTAAAGCAACCAACCGATGTCACTGCT 3370
 QY 3556 SerSerTrpThrGlyHisAsnCysSerArgGlySerArgThrGlyPhe 3571
 DB 3371 TCTTCTTGGACGGACATACATCTGTTCCAGAAAGAGGAGTGGTGT 3418

RESULT 22

US-09-984-429-525/c
 ; Sequence 525, Application US/09984429
 ; Publication No. US20040010132A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: 53 Human Secreted Proteins
 ; FILE REFERENCE: P2018P2
 ; CURRENT APPLICATION NUMBER: US/09/984,429
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,591
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 09/288,143
 ; PRIOR FILING DATE: 1999-04-08
 ; PRIOR APPLICATION NUMBER: PCT/US98/21142
 ; PRIOR FILING DATE: 1998-10-08

; PRIOR APPLICATION NUMBER: 60/061,463
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,529
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/071,498
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,527
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,536
 ; PRIOR FILING DATE: 1997-10-09
 ; PRIOR APPLICATION NUMBER: 60/061,532
 ; PRIOR FILING DATE: 1997-10-09
 ; NUMBER OF SEQ ID NOS: 727
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 525
 ; LENGTH: 7286
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-984-429-525
 Alignment Scores:
 Pred. No.: 0 Length: 7286
 Score: 5848.00 Matches: 1160
 Percent Similarity: 48.31% Conservative: 1
 Best Local Similarity: 48.27% Mismatches: 2
 Query Match: 29.28% Indels: 1242
 DB: 11 Gaps: 2
 US-09-977-053-4 (1-3571) x US-09-984-429-525 (1-7286)
 QY 2061 CysAsnAlaGlnGlySerTrpValProGluGlyGlnAspMetProArgCysIleAla 2080
 DB 7285 TGCATGCCAGGCAAGTGGGTACCCCAAGAGTCCAGACATGCCCGTTGTATAGCT 7226
 QY 2081 HisPheCysGluLysProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
 DB 7225 CATTTCTGTGAAACCTCCATCGGTTTCTATAGCATCTTGGAACTGTGAGCAAGCA 7166
 QY 2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
 DB 7165 AAATTTGCACTGGCTCAGTTGTGAGCTTTAAATGCAATGGAAGCTTTGTACTGAACACC 7106
 QY 2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
 DB 7105 TCAGCAAAAGATTGAATGATGAGAGGTGGGCAGTGGAGACCTTCCCCCATGTCCATCCAG 7046
 QY 2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
 DB 7045 TGCATCCCTGTGCGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCA 6996
 QY 2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
 DB 6985 AACTACAGTTTGTGAGCCATGGTGGCTTACAGCTGCCAACAAAGGGGTTCTACATCAAAAGGG 6926
 QY 2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerProIleProThrCysHis 2200
 DB 6925 GAAAGAAGAGCACCTGCCAAGACCAAGGCGAGTGGAGTAGTCTCTATACCGACGTGGCAC 6866
 QY 2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGlu----- 2216
 DB 6865 CCGGTATCTTGTGTGAAACCAACCTAAAGTTGAGAAATGGCTTTCTTGGAG-GGTAAGAGAGACCA 6807
 QY 2216 ----- 2216
 DB 6806 ATTAGCAATGGTCTGTGGCTTTTGTGCAGGGGCCCATATAATAATAAATAAACAACAA 6747
 QY 2216 ----- 2216
 DB 6746 ATAATTGAGTCCAAACACATGCCCAAGTTCATAAGGCATGGAACTGGAAATTACTATTAT 6687
 QY 2216 ----- 2216
 DB 6686 CACTAACAAAGAAATCATGATAATCTTTTAAACATTTTTCATAATAATTAATTCATCTCTTGATG 6627

QY 2216 ----- 2216
Db 6626 AATTAAATCATCTGCTAAGGTTGTGAATCACTTTCTGGAGGTAAGAGACTAAATTAGCAA 6567
QY 2216 ----- 2216
Db 6566 ATGCTCTGTGGCTTTCATCACTTACAGGGCCAGCCCACTGTGTAGTAAGCCACTATTTCAGA 6507
QY 2216 ----- 2216
Db 6506 AGCGGGCAATGAACAAATTAATTATTGAAAAACAGAAATGCTAAATCTACAGGTAGT 6447
QY 2216 ----- 2216
Db 6446 GAAGACTGGAAAAAATAATGAAGATGATTATGAATCTTTTGTAGTATATACTTTCTGCA 6387
QY 2216 ----- 2216
Db 6386 TTTTACTAGTATAGGCAATTTCAAGTGTCTTGTGGCAACAGCTTAGCCCTAAATAGGA 6327
QY 2216 ----- 2216
Db 6326 GTTAGAATAAACCAATTTTAATATAATATTTTGTTTTACAGAGGTTTTTTCAGGATGGATA 6267
QY 2216 ----- 2216
Db 6266 TATGCTTTATTTATAGAAGCAAGGTCTAAATGGAACTTAGAGAAATATATTATAATTT 6207
QY 2216 ----- 2216
Db 6206 TAAATGCTTAGACAAAAAGATAGGGTAAACCAATAATATGTAGTTTTTACAAATCTCTGTTTG 6147
QY 2216 ----- 2216
Db 6146 GTAAACCTAAAGGTAAATTTCACTGTGTATGACAAAGAGTTTCAGTCCAACCCAAATATGT 6087
QY 2216 ----- 2216
Db 6086 CCTGCTTTCAAGCTAAATGTGTCTTTCTCAATAAGCTTACCAAGAAATAGGAATGAGAA 6027
QY 2216 ----- 2216
Db 6026 TCTATTCGCCAAATCAAGGGCCATTTGMAATGACCATATTTCTAATTAACACATTTTTT 5967
QY 2216 ----- 2216
Db 5966 AAAACCCACATATATACACAGTACAAACCTTTTCTCCTTTTACATTCAAATAACAAAAACAG 5907
QY 2216 ----- 2216
Db 5906 TAAAAGCCTTTTATTTAAGGCTTTTGTCTAGATTTTAGTACTTTTTCATTGGAGAGGGTA 5847
QY 2216 ----- 2216
Db 5846 CAGAGTATAGTAGAGGTGTTCATTTATTAATTGGAAGAGGTATAGAGTATAAGTAG 5787
QY 2216 ----- 2216
Db 5786 AGAGGTGTTCAATTTATATAAATTTAATAAATATTTCTAGAGCTCTGCTTGGCACCAGGTAC 5727
QY 2216 ----- 2216
Db 5726 AGTTCTGAGAATCAGCGACTAGCATTTTAAACAAACACAGATTAGATGTGTAAATGGCATGGT 5667
QY 2216 ----- 2216
Db 5666 GCTTACATGATGGGAGACACAAAGCAATGCATAAACAAATGAGTAATTTTGTGTTATTT 5607
QY 2216 ----- 2216
Db 5606 TGTTTATGACCTGTTATATTACATGTTGTGAAATTTTTCAGATCACGAATATTAAATTTTTT 5547

QY 2216 ----- 2216
Db 5546 GTTTTAAAGATAGGATATATTATTTTCTATTTGTGTACAGCAGTAATTTTTTAAGAAATATGTT 5487
QY 2216 ----- 2216
Db 5486 AATATAAACAAAATTTTGTAGTAGTATTTCTTAAGACATACCCATAAAAAACATTTTATTGAC 5427
QY 2216 ----- 2216
Db 5426 TATTTATGATCTAACACAATTAACAATAATGTTAAATGTTTATGTGCTTCTTTACTAG 5367
QY 2216 ----- 2216
Db 5366 CATTTCTTTGGGCAGAAAGTTTTTTTGAACCTGAGTAATTTATCAATTTTAGCGTTGCCCTC 5307
QY 2216 ----- 2216
Db 5306 ATCAGTTCCACCAATCAATATATATAGATTACTGATCTTTTACTTGTATGGGCTCAAAACTG 5247
QY 2216 ----- 2216
Db 5246 TAGTAAGGATTTTCGACGTGGGATTTACTTAATTTGAATTTTCTTTTCTTTTCTTTTCT 5187
QY 2216 ----- 2216
Db 5186 TTCTTCCCTTTAGACAGAGTCTTGCTCTGTCTCCAGGCTGGAGTGCAGTGGCACAATCT 5127
QY 2216 ----- 2216
Db 5126 CAGCTCACTGCAACCTCTGCTCTCTGGGTTCAAGTGATTTCTCTGCTCAGCTTCCCAA 5067
QY 2216 ----- 2216
Db 5066 TGGCTGAGATTACAGGCACATACCACTGCTAGCTAATTTTGTACAGGTTTCACCAT 5007
QY 2216 ----- 2216
Db 5006 GTTGGCCAGGCTGTCTCGAACTCTCACTCAAGTGTCTCTGCTCTGCTCGGCTGCCAA 4947
QY 2216 ----- 2216
Db 4946 AGTGTCTGGATTGTAGGCATGAATCGTCATCGAGCCTTAAGTTGACTTTTCTACTATCAT 4887
QY 2216 ----- 2216
Db 4886 TTTCACTTATTTAAAAAATAGAAATGGATCTATTGGAAAAAACATAAATCATTTATTTGCT 4827
QY 2216 ----- 2216
Db 4826 TACTTCTTAATTTGATTCATTTTAAACATAGACCTTTTAGTTTTTTTCTACTATCCAGGATTT 4767
QY 2216 ----- 2216
Db 4766 TAGTTAATGCTATCATCTGTTATATACAAATGCGACTCACTGTCTTCTCTGTGTGCACAG 4707
QY 2217 HistHrThrGluValAlaGlyPheGluSerGluValAlaGlyGlnCysAsnProGlyTyrLys 2236
Db 4706 CATACACTGGCAGGATCTTTGAGAGTGAAGTGAATATCAGTGTAAACCGGCTATAAG 4647
QY 2237 SerValGlySerProValPheValCysGlnAlaAsnArgHisTyrPheHisSerGluSerPro 2256
Db 4646 TCGTCGGAGTCTCTGTATTTGTCTGCCAAGCCATCGCCACTTGGCAGTGAATCCCT 4587
QY 2257 LeuMetCysValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLys 2276
Db 4586 CTGATGTGTGTTCTCTCGACTGTGAAAAACCTCCCGATCCAGAAATGGCTTCATGAAA 4527
QY 2277 GlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeu 2296
Db 4526 GGAGAAAACTTTGAAGTAGGGTCCAAGGTTCAAGTTTCTGTAAATGAGGGTTATGAGCTT 4467
QY 2297 ValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyrPheAsnLysLysSerAsnPro 2316

Db 4466 GTTGGTACAGTTCTTGGACATGTCAGAAATCTGGCAATTCGAATTAAGAGTCAATCCA 4407
Qy LysCysMetProAlaLysCysProGluProProLeuLeuGlnLeuValLeuLys 2336
Db 4406 AAGTGCATGCTCCCAAGTGCAGAGCCGCCCTCTTGGAAACACGCTAGTATTAAAG 4347
Qy GluLeuThrThrGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGln 2356
Db 4346 GAGTTGACACCGAGGTAGAGTTGTGACATTTCTGTAAAGAGGCGCATGTCTTCGCA 4287
Qy GlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCys 2376
Db 4286 GGCCCTCTGCTCTGAAATGCTTGCCATCCACGAATGGAATGCACTCTTTCCCTGTTGT 4227
Qy LysIleValLeuCysThrProProProLeuIleSerPheGlyValProIleProSerSer 2396
Db 4226 AAGATTGTTCTTTGTACCCACCTCCCTAAATTTCTTGTGTGCTGCCCATATCTCTTCT 4167
Qy AlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGlyPhePheLeuAtqGly 2416
Db 4166 GCTCTTCATTTTGGAAAGTACTGTCAAGTATTTCTTGTGTAGGTGGGTTTTTCTTAAGAGGA 4107
Qy AsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysVal 2436
Db 4106 AATTCTACCACTCTCTGCCACCTGTATGGACCTCTCCACTGCCAGAAATGTGT 4047
Qy ProValGluCysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeu 2456
Db 4046 CCAGTAGAATGTCCCAACCTCGAGGAATCCCAATGGAATCAATTGATGTGCAAGGCCCT 3987
Qy AlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThr 2476
Db 3986 GCCTATCTCAGCACAGCTCTATACCTGACACCCAGGCTTTGAAATGTGTGGAAATACT 3927
Qy ThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyClyLysProThrCysLysAlaIle 2496
Db 3926 ACCACCTTTTGTGGAGAAATGCTCACTGCTTGGAGGAACCAACATGTAAGCCATT 3867
Qy GluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerThrAspLeuHisTyr 2516
Db 3866 GAGTGCCTGAAACCCCAAGGAGATTTTGAATGGCAAAATCTCTTACACGGACCTACACTAT 3807
Qy GlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeu 2536
Db 3806 GGACACGCTTACCTACTCTTGCAACCGAGGCTTTCCGCTCGAAGTCCAGTCCAGTCCCTG 3747
Qy ThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCys 2556
Db 3746 ACCTGTTTAGACAGAGTGAATGGATGTAGATGCCCATCTTGCATGCAATGCCACTGT 3687
Qy AspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerThrGlyAla 2576
Db 3686 GATTCCCCCAACCCCATTTGAAATGGTTTTTGTAGAAGGTGCAGATTACAGCTATGCTGCC 3627
Qy IleIleIleTyrSerCysPheProGlyPheGluValAlaGlyHisAlaMetGlnThrCys 2596
Db 3626 ATAAATCATCTACAGTTGCTTCCCTGGGTTCAGGTGGCTGTGATGCTCATGCGAGACCTGT 3567
Qy GluGluSerGlyTrpSerSerIleProThrCysMetProIleAspCysGlyLeuPro 2616
Db 3566 GAAGAGTCAGAGTGTCAAGTTCATCCCAACATGATATGCCATAGACTGTGGCTCCCT 3507
Qy ProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGln 2636
Db 3506 CCTCATATAGATTGTGAGACTGTACTAAACTCAAGATGACACCGAGGATATTTGAGCAA 3447
Qy GluAsnAspMetMetGluValProTyrValThrProHisProProTyrHisLeuGlyAla 2656
Db 3446 GAAGACACATGATGGAAGTTCCATATGTGACTCTCCTCACCTCTCTTATCATTTGGAGCA 3387
Qy ValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeu 2676

Db 3386 GTGGCTAAAAACCTGGGAAAAATACAAAGGAGTCTCTGCTACACATTTCATAAATCTCTG 3327
Qy TyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProVal 2696
Db 3326 TATGGTACCATTGGTTTCTACACTGTAACTCAGAGATATGAATCTTGGGGAACCTCTGTG 3267
Qy LeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGlu 2716
Db 3266 CTGATCTGCCAGGAGAGTGAATCTTGAATGGCAGTGCACCATCTCTGCAATTTCAATGAA 3207
Qy CysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGly 2736
Db 3206 TGTGACTGTGCTACTGCTCTCGAAATGGCTTTTGTGGTTTTTACAGAGACTAGCATGGGA 3147
Qy SerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeu 2756
Db 3146 AGTGTGTGAGATATAGCTGTAAACCTTGGACACATTTCTAGCAGGCTCTGACTTAAGGCTT 3087
Qy CysLeuGluAsnArgLysTrpSerGlyValAserProArgCysGluAlaIleSerCysLys 2776
Db 3086 TGTCTAGAGAAATAGAAAGTGGAGTGGTGCCTCCCAACGCTGTGAAGCCATTTTCATGCAA 3027
Qy LysProAsnProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThr 2796
Db 3026 AAGCCAAATCCAGTCAATGGAATGGAATCATCAAGGAAGCACTACACATACCTGAGCAG 2967
Qy LeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGln 2816
Db 2966 TTGTACTATCAGTGTGACCCCGGATATGTGCTGAATGGCACTGAGAGGAGAAACATGCCAG 2907
Qy AspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerPro 2836
Db 2906 GATGACAAAACTGGGATGAGGATGAGCCCATTTGCAATCTCTGTGCACTGCGATTCCACC 2847
Qy ProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGlu 2856
Db 2846 CCAGTCTCAGCCCAATGGCCAGGTGAGAGAGACGAGTACACATTTCCAAAAAGAGATTGAA 2787
Qy TyrThrCysAsnGluGlyPheLeuGluIleAlaArgSerArgValCysLeuAlaAsn 2876
Db 2786 TACATTTGCATGAAGGGTCTTGTCTTGAGGAGGAGGAGTGGGTTTGTCTTGCCAA 2727
Qy GlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProGln 2896
Db 2726 GGAAGTTGAGTGGAGCCACTCCCGACTGTGCTGCTGTGCAATGTGCCACCCGCCAA 2667
Qy LeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHis 2916
Db 2666 CTGGCCAAATGGGGTGACGGAAGGCTTGGACTATGGCTTCATGAAGGAAGTAACATTCCAC 2607
Qy CysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsn 2936
Db 2606 TGTACAGAGGCTACATCTTGCACGGTGTCTCAAACTCAGTGTGAGAGTCTTCCAGAGTCC 2547
Qy TrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLeuAla 2956
Db 2546 TGGATGACAGAGATTCTCTCTGTAAACCACTCACTGTGAGCCCTCTCTGAAGATCTGCC 2487
Qy HisGlyPheProAsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPhe 2976
Db 2486 CATGGTTTCCCTAAATGGTTTTTCTTTATTCATGGGGCCATATACAGTATCAGTGTCTT 2427
Qy ProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSer 2996
Db 2426 CCGTGTATAGCTCCATGGAATTCATCAAGAGGTGCTCTCTCCATGGCTCTCTGGAGT 2367
Qy GlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThr 3016
Db 2366 GGCAGCTCACCTCTCTGCTGCTTGCAGATGTTCCACACAGTAAATGAATATGGAAT 2307
Qy ValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPhe 3036
Db 2306 GTCAATGGGACAGATTTTGTGCTGGAAAGGACGCCCGGATTCAGTGTCTTCAAGGCTTC 2247

Qy	3037	LysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPhe	3056	Qy	3146	---	---	---	3146
Db	2246	ARGCTCTAGGACCTTCTGAAATCACCTGTGAGCGCATGCGCCAGTGGAGCTCTGGGTTC	2187	Db	1167	TGTATTCTTACTAGTATCTGAGTCCATGTAGGTATTTCATAATACATATGAATATGCTGT	1108	Db	3146
Qy	3057	ProHisCysGluHisThrSerCysGlySerLeuProMetIleProhenAlaPheIleSer	3076	Qy	3146	---	---	---	3146
Db	2186	CCCCACTGTGAACACACTTCTTGTGGTCTCTTCCAAATGATACAAATGGTTCATCAGT	2127	Db	1107	ATTCTTACTAGTATCTGAGTCCATGTAGGTATTTCATAATACATATGAATTTATGCTGTATT	1048	Db	3146
Qy	3077	GluThrSerSerTrpLysGluIleValIleThrTrpSerCysArgSerGlyTrpValIle	3096	Qy	3146	---	---	---	3146
Db	2126	GAGACCAAGCTCTTGGAAAGGAAATGTGATAACTTACAGCTGCAGGCTGGATATGCATA	2067	Db	1047	CTTACTAGTATCTGAGTCCATGTAGGTATTTCATAATACATATTAATTATGCTGTATTCT	988	Db	3146
Qy	3097	GlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpProVal	3116	Qy	3146	---	---	---	3146
Db	2066	CAAGCCAGTTCAGATCTGATTTGTACAGAGAAAGGGGTATGAGCCAGCCCTTATCCAGTC	2007	Db	987	ACTAGTATCTGAGTCCATTTTGGACTGTCTAGAAATTTCTGAACTGTGTAGCTGTTTATAG	928	Db	3146
Qy	3117	CysGluProLeuSerCysGlySerProProSerValAlaIleValAlaThrGlyGlu	3136	Qy	3146	---	---	---	3146
Db	2006	TGTGAGCCCTTGTCTGTGGTCCCAACCGTCTGTCCCAATGCACTGGCACTGGAGAG	1947	Db	867	TAAAGTAAATTTTAGTCTTGTAACTTTTTCATGTGGTAATCTGAGAACTGATCTGGTATT	808	Db	3146
Qy	3137	AlaHisThrTrpGluSerGluValLysLeu	3146	Qy	3146	---	---	---	3146
Db	1946	GCACACACCTATGAAGTGAAGTGAAGTCT - CAGTAAGATCATGGCTCTAGGGTCTCCAA	1888	Db	807	TACCGGTTTTTGTGTTTTTGTGTTTTTGTGAGACAAGCTCTTGTCTGTCTCACCAGGCTGGA	748	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1887	GGGTGCGGCGGCGGNTATGAAATGAATATGTATGATCAAAATTTTGTGATGGAA	1828	Db	747	GTGAGTGGCATGTTCTTGGTCACTGAGCCCTCAACCTCTTGTAGTTCAAGCAGTCTCTCC	688	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1827	GAAAAACATCATCTCTGAACTGCTCTCCAAACACCTGGTGCTTTTAAACACTTGGCTTTTATTC	1768	Db	687	CACCTCTGCTCCCAAGTAGCTGAGACTACAGCATGTGCCACGTGCCACGTAAATTTT	628	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1767	TACCTACAATATGTAGAAATCCACTGAGTATATAAAATNAATCGATACTTTTTCAGACA	1708	Db	627	GTTTATTTTTTATAGATGAAGTCTCACTATGTTTCCAGCGCTAGTCTCAAGCTCTCTGG	568	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1707	TGTTCTTAATATATATCATGACTATCATATCTCTGTTGAGCTTCGTGAAACAAACAGCA	1648	Db	567	ACTCAAGTGTCTCTCTGCTCAGCCTTCCAAAGTGCAGGGATTAAGGATGAGCCACC	508	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1647	AGATCCTCTAGCTAGTTGAAAGCTCCGGAATCGTGAAGCTGTTTTTAAACATCAGGA	1588	Db	507	GCACCCAGCTCAACTGGTGTATTGTTATGTTGGTGTGTTTTCATGGCAATGCCATCG	448	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1587	TAGTAAATTTCTTTATGATTATAACTTTTGTGATTATATTGACACATGCAATATTTGGGA	1528	Db	447	CTGACATATTGATTGTGCAACAATCAATATGTTGCACAAAATATCTACTGAGCCTTGTA	388	Db	3146
Qy	3146	---	3146	Qy	3146	---	---	---	3146
Db	1527	CATAAGGCAAAAACTGAGAAATTCAGGCCATGATCATGCCAATATCATAGGATTAAG	1468	Db	387	ACAGCTCTTAAATCCAAAGTAGATCAGAAACTATTTGTACTTCAATTAAGGCACTCA	328	Db	3146
Qy	3146	---	3146	Qy	3147	---	---	---	3146
Db	1467	AAATATCTTATGTAAGGACCTAGAGCAGCACTTGGCACAAGTAGGTATTTCATAATACA	1408	Qy	327	TTTTTGCAATTTCCATGTCAGATGTCTGGAGGTATACGATGGATACAGATACAGATA	268	Qy	3160
Qy	3146	---	3146	Db	3161	PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys	3180	Qy	208
Db	1407	TATGAATTTATGCTGTATTCCTACTAGTATCTGAGTCCATGTAGGTATTTCATAATACATAT	1348	Qy	267	TTCCCTGTGAGAAAGTGTGGTCTGGTCCCTGAGAGAAATCTCTCTGAGTCTCTAAAAA	208	Qy	3200
Qy	3146	---	3146	Db	3181	CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspPheSerValAsn	3200	Qy	148
Db	1347	GAATTTATGCTGTATTCCTACTAGTATCTGAGTCCACGTAGTATTCATAATACATATTA	1288	Qy	207	TGTCCTCTCCCGGAAACATAACACATATACCTTTGTATCATGCGGACCATTTTCACTGTGAAAT	148	Qy	3201
Qy	3146	---	3146	Db	3201	ArgGlnValSerValSerCysAlaGluGlyTrpThrPheGluGlyValAsnIleSerVal	3220	Qy	88
Db	1287	TTATGCTGTATTCCTACTAGTATCTGAGTCCACGTAGTATTCATAATACATATTA	1228	Db	147	AGGCAAGTTCTGTGTCATGTCAGAGAGGGTATACCTTTGAGGGAGTTAACATATCAGTA	88	Qy	3221
Qy	3146	---	3146	Qy	3221	CysGlnLeu	3223	Qy	
Db	1227	TGCTGTATTCCTACTAGTATCTGAGTCCACGTAGTATTCATAATACAAATGAATATGTC	1168	Db					

```
Db      87  TGTGAGGTA 79
|||||...
; Sequence 34, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803 1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 34
; LENGTH: 3448
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (415)..(2991)
US-10-302-172-34

Alignment Scores:
Pred. No.: 0 Length: 3448
Score: 5485.00 Matches: 982
Percent Similarity: 99.90% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 27,46% Indels: 0
DB: 13 Gaps: 0

US-09-977-053-4 (1-3571) x US-10-302-172-34 (1-3448)

QY 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyLe 337
Db 1 TATGATGCAAGCTTCCCTCGGGGACATACAACTGAGCCTCACCAGGAGGATC 60
QY 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
Db 61 AGCAGTTGCATTCATGTCCTCCGATGAAATCACCCTCCACCTGGAAGCAGCATCCCT 120
QY 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
Db 121 GAAGACTGTGTCGAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGCCAC 180
QY 378 CysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnHis 397
Db 181 TGCCCTGACCTGAGCCTCCGAAATGTTACTTATCCAAACACTTGCAACACACCAC 240
QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerIle 417
Db 241 TTCAATGAGCCCTGTGGGGTCGATGTCACCTCCGATTTGATTTGTTGGGAGAGCAGCATC 300
QY 418 IleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThr 437
Db 301 ATCTATGTCTACCAATGGTTTGGTCCGGTTACAGAGCTTACTGAGAGTAAGAACCA 360
QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
Db 361 TGTCTCATCTCCGACCGGAAACATGCGCCACATCAGCTGTTCTACAGGGAAATGTTA 420
QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
|||||...

421 TATAAGACAAACATGTTTGGTCTGTGATCAAGGGTACAGACTAGAAGCGCAGTATAAG 480
QY 478 LeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHis 497
Db 481 CTTACTTTGTCAAGAAACAGCCAGTGGGATGGGCGAAGCCCGGTGTGTGGAGCGCCAC 540
QY 498 CysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGln 517
Db 541 TGTTCACCTTTTCAGATGCCCAAGATGTTCATCATATCCCCCACAACATGTGGCAAGCAG 600
QY 518 ProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGly 537
Db 601 CCAGCCAAATTTGGGACGATCTGTATGTAAAGTTGCGCGCAAGGGTTCATTTTATCTGGA 660
QY 538 ValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAla 557
Db 661 GTCAAGAAATGCTGAGATGTACCATCTTCGAAAAATGGAAATGTCGGAGTTCAGGCGACT 720
QY 558 ValCysLysAspValGluAlaProGlnIleAsnProLysAspIleGluAlaLysThr 577
Db 721 GTGTGTAAAGACGTGGAGGCTCTCAAAATCACTGTCTTAAGGACATAGAGGCTTAAGACT 780
QY 578 LeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSer 597
Db 781 CTGAAACAGCAGATCTGCAATGTTCCTGGCAGATTCACACAGCTAAAGACAACCTCT 840
QY 598 GlyLysValSerValHisValHisProAlaPheThrProTyrLeuPheProIle 617
Db 841 GGTGAAAGGTGTCAGTCCACGTTCATCCAGCTTTCACCCACCTTACCTTTTCCCAAT 900
QY 618 GlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
Db 901 GGAGATGTTGCTATCGTATACACGCAACTGACTATCCGCAACAGGCGCAGCTGCATT 960
QY 638 PheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProPro 657
Db 961 TTCATATCAAGGTTATTTGATGAGAACACCACTGTCTATAGACTGGTGACAGTCTCCACCT 1020
QY 658 ProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAsp 677
Db 1021 CCCGTCAGGCTCTCGAGAGAGGTACATGCGCGAGCTGGGATGAGCTCTCAGTTCTCAGAC 1080
QY 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697
Db 1081 AACTCAGGGGCTGAAATTTGGTCAATTTACAGAGTACATACAGAGGAGACCTTTTCCCTCAA 1140
QY 698 GlyLysThrIleValGlnTyrThrAlaThrAspSerGlyAsnAsnArgThrCysAsp 717
Db 1141 GGGGAGACTATAGTACAGTATACAGCCACTGACCCCTCCAGGCAATAACAGGACATGTGAT 1200
QY 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
Db 1201 ATCCATATTGTCATAAAGGTTCTCCCTGTGAATTTCAATTCACCTGTAAATGGGAT 1260
QY 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
Db 1261 TTTATATGACCTCCAGATAATCTCGGAGTCAACTGTACATTAACCTGCTGGAGGGCTAT 1320
QY 758 AspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLys 777
Db 1321 GATTTTCAGAGAGGGTCTACTGACAAAGTATTTATGTGCTTATGAAGATGGCGTCTGAAA 1380
QY 778 ProThrTyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPhe 797
Db 1381 CCACATATACCACTGATGAGGCGAGCTGTGCCAAAAAAGCTTTTGCAAAACCGCGGTTTC 1440
QY 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLys 817
Db 1441 AAGTCTCTTGGAGATGTTCTACAAAGCAGCTCGTTGTGTGATGACACAGATCTGATGAAGAAG 1500
QY 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
Db 1501 TTTTCTGAAGCATTTGAGACGACCCCTGGGAAAAAATGGTCCCATCATTTTGTAGTATGCA 1560
```

QY 838 GluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsn 857
DB 1561 GAGGACATTCAGCTGAGAGGAGACCTGACCAAAAAATATTCCTAGAAATATAAT 1620
QY 858 TyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTgGlyValaAlaAsnArgLeu 877
DB 1621 TATGACTATGAAATGGCTTTTGCATTCGACCAAGTGGCTGGGGTGAGCTATAGGCTG 1680
QY 878 AspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsn 897
DB 1681 GATTACTCTTACGATGACTCTCTGACACTGTGCAAGAAACAGCCACAAGCATGCGCAAT 1740
QY 898 AlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIle 917
DB 1741 GCCAGTCTCTCAGCGATTAAGAAAGTGCCTCATTTATCTGACTATATAAATTAAGTTAAT 1800
QY 918 PheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGlu 937
DB 1801 TTTAACATCACAGCTAGTGTGCTATTCCTGATGAAAGAAATGATACCTTGAAATGGGA 1860
QY 938 AsnGlnGlnArgLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeu 957
DB 1861 AATCAGCAACGACTCTTTCAGACTTGGAACTATCACAAATAAATGAAAAGGACTCTC 1920
QY 958 AsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsn 977
DB 1921 AACAAAGACCCCATGATTTCTTTCAGCTTGATCAGAAATATCTTATAGCCGACAGCAAT 1980
QY 978 SerLeuGluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArg 997
DB 1981 TCATTAGAACAACAAAGAGCTTCCCTCTCTGACACAGGCTCAGTCTGAGAGGGGT 2040
QY 998 MetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer 1017
DB 2041 ATGTGTGTCATTTGCGTCTTGGAACTATTATATCTGGAACATTTACCTGTGAAAGC 2100
QY 1018 CysArgIleGlySerTyrGlnAspGluGlnLeuGluCysLysLeuCysProSer 1037
DB 2101 TGCCGATCGGATCCTATCAGATGAGAGGCACTTGAGTGCAGCTTTGCCCTCT 2160
QY 1038 GlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLys 1057
DB 2161 GGGATGTACAGGAATATATCCTTCAAGAAACATCTCTGATTGTAAAGCTCAGTGTAAA 2220
QY 1058 GlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyr 1077
DB 2221 CAAGGCACCTACTCATCAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCATTTAT 2280
QY 1078 GlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLys 1097
DB 2281 CAGCCAAAATTTGGTTCCCGAGCTGCTCTCGTGTCCAGAAACACCTCACTGTGAAA 2340
QY 1098 ArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArg 1117
DB 2341 AGAGGAGCGGTGAACATTTCTGCATGTGGATTCCTTGTCCAGAGGAAATTTCTGGGT 2400
QY 1118 SerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLys 1137
DB 2401 TCTGGTTAATGCCCTGTCAACCATGCTCTCGTGAATATACCAACCTTAATGAGGGAAG 2460
QY 1138 AlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIle 1157
DB 2461 GCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
QY 1158 ThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSerValValProPro 1177
DB 2521 ACAGATGTTTCAAGTTTATTTTCACTTCTTCTCAGCGGACAGAGAAAGTGTGTCCTCT 2580
QY 1178 AlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCys 1197
DB 2581 GCCTCTCTTGGACATATTAAGAGGATGAATCAGCAGTTCAGGTTTTCATGATGC 2640

QY 1198 PhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGluArgGlyTyrValCys 1217
DB 2641 TTCTTTAAACCTTGTCCACATAGTAGGAACCTGCCAGCAACTTGGCGCTGTTATGTTGT 2700
QY 1218 LeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerPro 1237
DB 2701 CTCTGTCCACTTGGATATACAGGCTTAAAGTGTGAAACAGACATCATGATGAGTGCAGCCCA 2760
QY 1238 LeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCys 1257
DB 2761 CTGCTTCTGCTCAACATAGGAGTTTGTAAAGACCTAGTTTGGGGATTCATTTGTGAGTGC 2820
QY 1258 ProSerGlyTyrThrGlyGlnArgCysGluGlnAsnIleAsnGluCysSerSerPro 1277
DB 2821 CCATCAGGTTCACAGGTCAGCGGTGTGAAGAAATATTAATAGTGTAGTCCAGTCT 2880
QY 1278 CysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLys 1297
DB 2881 TGTTTAAATAAAGGAATCTGTGTGATGCTGGCTGCTATCGTTGCACATGTGTGAAA 2940
QY 1298 GlyPheValGly 1301
DB 2941 GGAATTTGTAGT 2952

RESULT 24

US-10-466-164-29
; Sequence 29, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFFO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN, Jr., Edward H.;
; APPLICANT: PERALTA, Careyna H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT
; CURRENT APPLICATION NUMBER: US/10/466,164
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 3262
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 LI:202943.4:2001JAN12
US-10-466-164-29

Alignment Scores:

Pred. No.:	0	Length:	3262
Score:	4930.00	Matches:	947
Percent Similarity:	95.33%	Conservative:	12
Best Local Similarity:	94.14%	Mismatches:	20
Query Match:	24.68%	Indels:	28
DB:	13	Gaps:	4

US-09-977-053-4 (1-3571) x US-10-466-164-29 (1-3262)

QY	1590	LysSerLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpPro	1609
DB	289	GAGTCACTGGCTACCTCCCTCCAGAGAACTCAGTAAAGAAACGGTTAGCATGGCCT	348
QY	1610	AspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysSerIlePheCysSer	1629
DB	349	GATTCTTGTGAGGAATTTGGGGAAAGTGAAGATCGATTCTAAGAGCATATTGTGTCT	408
QY	1630	AspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLys	1649
DB	409	GGTTGCCACGCTTAGGAGGCTCAGTGCCTCATCTGAGAACTGCATCTGAAGATTAAAG	468
QY	1650	ProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProVal	1669
DB	469	CCAGGTTCCAAAGTCAATCTGTTCTGTGATCCAGGCTTCCAGCTGGTCGGGAACCTGTG	528
QY	1670	GlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSer	1689
DB	529	CAGTACTGTCTGANTCAGGACAGTGGACAAACCACTTCTCCTGCTGTAACGCTTAGC	588
QY	1690	CysGlyValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGly	1709
DB	589	TGTGGGTCGCCACTCTCTTGGAGATGGCTTCCATTTCCAGCCGATGACTTCTATGCTGGC	648
QY	1710	SerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe	1729
DB	649	AGCACAGTAACTACCAAGTGCACCAATGGCTACTATCTATTGGGTGACTCAGGGATGTT	708
QY	1730	CysThrAspAsnGlySer-TripAsnGlyValSerProSerCysLeuAspValAspGluC	1749
DB	709	TGTACAGATATGGGAGCTAGGAACGGGCTTCCACCATCTCCGCTGATGTCATGATG	768
QY	1749	AlaValGlySerAspCysSerGluHisAlaSerCysLeu-AsnValAspGlySerTyrI	1769
DB	769	TGCAGTTGGATCAGATTGATGAGCATGCTCTTCTGCTGAAACGTAGATGGATCTTACA	828
QY	1769	IeCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysC	1789
DB	829	TATGTTTCATGTGTCCACCGTACACAGGAGATGGGAAACCTGTGCAGAACCTTATAAAT	888
QY	1789	ys-LysAlaProGly--AsnProGluAsnGlyHisSerSerGlyGluIleTyrThr-Va	1807
DB	889	GCTAAGGCTCCAGCGCAGATCCGGAATATGGCCACTCTCTCAGTGAGATTATACAGT	948
QY	1807	lGlyAla---AlaValThrPheSerCysGlnGluGlyTyr-GlnLeuMetGlyValThrL	1826
DB	949	AGGTGCCGGAAGTCACATTATTACGTGTGAGGAAGATACCCAGTTGATGGGAGTAAACA	1008
QY	1826	ylleThrCysLeuGlu-SerGlyGluTrpAsnHisLeuIlePro-TyrCysLysAlaVa	1845
DB	1009	AAATACATGTTTGGAGTACTCGAGAAATGGAATCATCTAATACCAATATTGTAAAGCTGT	1068
QY	1845	lSerCysGlyLysPro-AlaIleProGluAsnGlyCysIleGlu-GluLeuAla---Phe	1863
DB	1069	TTTANTGTGTAAACCGGACTATTCAGAAATATGGTTGCATTGAGCGGATAGCCACTTTT	1128
QY	1864	ThrPhe---GlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAsp	1882
DB	1129	ACCTATTGGCGAGCAAGTGCATATAGTGTGTAATAAAGGATATATCTCTGGCCGGTAT	1188
QY	1883	LysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCysGluPro	1902
DB	1189	AAAGAATCATCTGCTTCTGCTAACAGTCTTGTGAGTCAATTCCTCTCTGTTGTGAACCA	1248

QY	1903	ValLysCysSerSerProGluAsnIleAsnAsnGlyLys-TyrIleLeuSerGlyLeuTh	1922
DB	1249	GTCAAGTGTCTAGTCCGGAACAATATAAATAAATGACAACTATATATAGAGTGGGCTTAC	1308
QY	1922	rTyrLeuSerThrAlaSer-TyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer-	1941
DB	1309	CTACCTTTCTACTGTCATATAATTCATGCGCATACAGGATACAGCTTACAGGGCCCTTCCC	1368
QY	1942	IleIleGluCysThrAlaSerGlyIleTrpAspArgAlaPro-ProAlaCysHisLeuVa	1961
DB	1369	ATTATTGAAATGACGGCTTCTGGCAATCTGGACAGAGCGGCACCTGCTGTCACTGT	1428
QY	1961	lPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPh	1981
DB	1429	CTTCTGTGGAGAACCACTGCTCCATCAAAAGATGCTGTCTATTACGGGAATAAATTCAC	1488
QY	1981	eArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrI	2001
DB	1489	CAGAACACCGCTCCTTACACTTTCACAAAGAGGCTATATCTCTTGTGCTTGTACACCAT	1548
QY	2001	eGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCy	2021
DB	1549	TGAATGCTTGCCGACGCGCAAGTGTAGTAGAGTACAGCGTGCCTGGCTGTCTCTG	1608
QY	2021	sAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAs	2041
DB	1609	TGATGAGCCACCATTTGTGGACCAAGCTCTTCCAGAGACTGCCCATCGCTCTTCGGAGA	1668
QY	2041	lIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCy	2061
DB	1669	CATTGCAATCTACTACTCTCTGATGTTTACAGCTTAGCAGACAATTTCCCACTTCTCTG	1728
QY	2061	sAsnAlaGlnGlyLysTrp-ValProProGluGlyGlnAspMetProArgCysIleAlaH	2081
DB	1729	CAATGCCAGGCGCAAGTGGGGTACCCCAAGAGTCAAGCATGCCCGCTGTGTATAGCTC	1788
QY	2081	isPheCysGlu-LysProProSerValSerTyrSerIleLeuGluSerValSerLysAla	2100
DB	1789	ATTCTGTGAAAAAACCTCCATCGTTTCTCTATAGCATCTTGGAACTCTGTGAGCAAGCA	1848
QY	2101	LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr	2120
DB	1849	AAATTTGCACTGGCTCAGTTGTGAGCTTTAAATGCAATGGAAGGCTTTGTACTGAACAC	1908
QY	2121	SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln	2140
DB	1909	TCAGCAAGATTTGATATGAGAGTGGGAGTGGGAGACCTTCCCCCATGTCCATCCAG	1968
QY	2141	CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer	2160
DB	1969	TGCATCCCTGTGCGGTGTGGAGAGCCACCAAGCATCATGAATGCTATGCAAGTGGATCA	2028
QY	2161	AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly	2180
DB	2029	AACTACAGTTTTTGGAGCCATGGTGGCTTACAGCTGCAACAGGGGTTCTACATCAAGGG	2088
QY	2181	GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis	2200
DB	2089	GAAGAAGAGCACCTGCGAAGCCACAGGCGAGTGGAGTAGTCTCTTACCGAGCTGCCAC	2148
QY	2201	ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrGly	2220
DB	2149	CCGGTATCTTGTGTGTAACCCACCTAAGTTTGAGAAATGGCTTTCTGGAGCATACAACTGC	2208
QY	2221	ArgIlePheGluSerGluVal-ArgTyrGlnCysAsnProGlyTyrLysSerValGlySe	2240
DB	2209	AGGATCTTTGAGAGTGAAGTTGAGGTATCAGTGTAAACCGCGCTTATAAGTCAGTCGGAAG	2268
QY	2240	rProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVa	2260
DB	2269	TCTGTATTGTCTGCGACGCAATGCCACTGGCACAGTGAATCCCTCTGTATGTGTGT	2328
QY	2260	lProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPh	2280

```

Db 2329 TCCTCTCGACTGTGGAACACCTCCCGCATCAGAAATGGCTTCATGAAAGGAGAAACCTT 2388
Qy eGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSe 2300
Db TGAAGTAGGGTCCAGGTTTCAGTTTCTCTGTANTGAGGGTATGAGCTGTGTGGTGACAG 2448
Qy rSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPr 2320
Db TTCTTGACATGTCAGAAATCTGGCAATCTGCAATGGAATGGAATGCAATCCAAAGTGCATGCC 2508
Qy oAlaLysCysProGluProPheLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrTh 2340
Db TGCCAAAGTCCAGAGCGCCCTCTGGAACACAGCTAGTATTAAAGGAGTTGACCAC 2568
Qy rGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVa 2360
Db CGAGTAGGAGTGTGACATTTCTCTGTAAGAGGCGCATGTCTTGCAAGGCCCTCTCTGT 2628
Qy lLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysLysIleValLe 2380
Db CCTGAAATGCTTGCATCCAGCAATGGAATGACTCTTTCCCTGTGTGTAAGATTTGTTCT 2688
Qy uCysThrProProPheLeuIleSerPheGlyValProIleProSerSerAlaLeuHisph 2400
Db TTGTACCCACCTCCCTTAAATTTCTTTGTGTGTCCTCCCATTTCTTCTGCTCTTCATTT 2748
Qy eGlySerThrValLys-TyrSer-CysValGlyGlyPhePheLeu-ArgGlyAsnSerTh 2419
Db TGGAGTAGTCTGTCAAGTATTTCTTGATGTAGTGGGTTTCTTAAGCAGGAAATTTCTAC 2808
Qy rThrLeuCysGlnProAspGlyThrTrp-SerSerProLeuProGluCysValProValG 2439
Db CACCTCTGCCAACCTGATGCACTGGAGGCTCTCCACTGACAGAAATGTGTTCAGTAG 2868
Qy luCysProGlnProGluGluIleProAsnGlyLysIleLeuValGlnGlyLeuAlaTyrL 2459
Db AATGTCTCCCACTCAGGAAATCTCCCAATGGAAATCAATGATGTGCAAGGCCCTTGCTATC 2928
Qy euSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrL 2479
Db TCAGCAGAGCTCTCTATCTGTCAGACCGAGGCTTGTGAATGTGGGAAATCTACCACCT 2988
Qy euCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCys- 2498
Db TTGTGGAGAAATGCTCACTGGCTTGGAGGAGAAACCAAGATGTAAAGCCATTGAGTGCC 3048
Qy LeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGln 2518
Db GTGAACCCCAAGAGATTTGAATGGCAAAATTTCTTACCGGACCTACACTATGAGCAG 3108
Qy ThrValThrTyrSerCysAsn-ArgGlyPheArgLeuGluGlyProSerAlaLeuThrCy 2538
Db ACCGTACTACTCTTGCACACCGAGGCTTTCGGCTGGAAGTCCCGAGTGCCTTGACCTG 3168
Qy sLeuGluThrGlyAspTrpAspValAsp-AlaProSerCysAsnAlaIleHisCysAspS 2558
Db TTTTAGAGACAGGTGATGGGATGTAGATTTGCCCATCTTTGCAATAGCATCCACTGTGATT 3228
Qy 2558 exProGlnProIleGluAsnGlyPheValGlu 2568
Db 3229 CCCCACCAACCATTTGA--AATGGTTTTGTAAAG 3258

```

RESULT 25

US-09-728-952-93

; Sequence 93, Application US/09728952

; Patent No. US20020111302A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

```

; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Vicki
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 799
; CURRENT APPLICATION NUMBER: US/09/728,952
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 93
; LENGTH: 2437
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (184)..(2172)
US-09-728-952-93

Alignment Scores:
Pred. No.: 0 Length: 2437
Score: 4058.00 Matches: 716
Percent Similarity: 99.17% Conservative: 1
Best Local Similarity: 99.03% Mismatches: 6
Query Match: 20.32% Indels: 1
DB: 9 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-728-952-93 (1-2437)

Qy 2849 TyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAla 2868
Db 2 TACACATTTCCAAAAGAGATTGA-TACACTTGCATGAAGAGGGTTCTTGCTTAGAGGAGCC 60
Qy 2869 ArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValPro 2888
Db 61 AGGAGTGGGTTTGCTTGGCAATGGAACTTGGAGTGGAGGCCACTCCCGACTGTGTGCT 120
Qy 2889 ValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGly 2908
Db 121 GTCAGATGTGCCACCCGCCACAACTGGCCAAATGGGCTGACCGAAGGCCCTGGACTATGGC 180
Qy 2909 PheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProLys 2928
Db 181 TTCATGAAGAGAGTAACTTCCACTGTCAAGGGCTACATCTTGACGGTGTCTCCAAA 240
Qy 2929 LeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsn 2948
Db 241 CTCACCTGTTCAGTCAGATGGCAACTGGGATGGAGATTCCTCTCTGTAACCCAGTCAAC 300
Qy 2949 CysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGly 2968
Db 301 TGTGGACCTCTCGAAGATCTTGCCCATGGTTTCCCTTAATGGTTTTTCTTTATTCATGGG 360
Qy 2969 GlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAsnSerSerArgArg 2988
Db 361 GGCATATACAGTATACAGTCTTCTGTTATAGCTTCATGGAAATTCATCAGAGAGG 420
Qy 2989 CysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSer 3008
Db 421 TGCCTCTCCAAATGGCTCTCGAGTGGCAGCTACCTCTCTGCTGCTGCTGCGATGTTCC 480
Qy 3009 ThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAla 3028
Db 481 ACACCAAGTAAATTGAATATGAACTGTCAATGGACAGATTTTGACTGTGGAAAGGAGGCC 540
Qy 3029 ArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAla 3048
Db 541 CGGATTCAGTCTTCGAAGCTTCAAGCTTCAAGCTTCTAGGACTTCTGAAATCACCTGTGAGCC 600
Qy 3049 AspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuPro 3068

```


Db 601 GATGCCAGTCGAGCTCTGGGTTCACCACTTTGAACACACACTTCTTGTTGCTCTTCCA 660
Qy
Qy 3069 MetIleProAsnAlaPheIleSerGluThrSerSerTrpIysGluAsnValIleThrTyr 3088
Db 661 ATGATACCAAAATCGCTTCATCAGTGGAGACAGCTCTTGGAAAGGAAATGTGTAACCTTAC 720
Qy 3089 SerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGly 3108
Db 721 AGCTGCAGTCTGGATATGTATACAGGCGAGTTCAGATCTGATTTGTACAGAGAAGGG 780
Qy 3109 ValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerVal 3128
Db 781 GTATGAGGCGAGCTTATCCAGTCTGTGAGCCCTTCTCTGGTGGTCCCGACCGTCTGTG 840
Qy 3129 AlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValIysLeuArgCys 3148
Db 841 GCCAAATGCAATGGCAATCGAGAGGACACACCTATGAAGTGAAGTGAACACTCAGATGT 900
Qy 3149 LeuGluGlyTyrThrMetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArg 3168
Db 901 CTGGAAGGTATTACGATGATACAGATACAGATCAATCACTGTCCAGAAAGATGTGCG 960
Qy 3169 TrpPheProGluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleThr 3188
Db 961 TGGTTCCTCGAGAGATCTCTGCACTCTCTAAATAATGCTCTCTCCCGGAAAAACAATA 1020
Qy 3189 HisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAla 3208
Db 1021 CATATACTTGTATCATCGGAGCGATTTCACTGTGTGAATAGGCAAGTCTCTGTGTCTATG 1080
Qy 3209 GluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGlu 3228
Db 1081 GAAGGTATACCTTTGAGGGAGTTAATCATATCACTATGTCAGCTTGTGGAACCTGGAG 1140
Qy 3229 ProProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGlu 3248
Db 1141 CCACCAATCTCCGATGAAATCTGCACTCCAGTTCTTGTGGGAAACCTGAAAGTCCAGAA 1200
Qy 3249 HisGlyPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGlu 3268
Db 1201 CATGATTTGTGTGGCAGTAAATACCTTTGAAGACAAATTTATTCAGTGTGAG 1260
Qy 3269 ProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSer 3288
Db 1261 CTGTGCTATGAATCAGAGGGAGACAGGAAAGTCTGTGCCAGGAGAACAGACAGTGGAGT 1320
Qy 3289 GlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGly 3308
Db 1321 GGAGGGGTGGCAATATGCAAGAGACACAGGTGTGAAACTCCACTTGAAATTTCTCAATGG 1380
Qy 3309 LysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGly 3328
Db 1381 AAAGCTGACATTGAAACACAGGACGATGACACCACTGGTATATTTCTGCAACAGAGGC 1440
Qy 3329 TyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisPro 3348
Db 1441 TACAGTCTTGAGGGCCATCTGAGGACACACTGACAGAAATGGAACTGGAGCCACCCA 1500
Qy 3349 ValProLeuCysGlyProAsnProCysProValProPheValIleProGluAsnAlaLeu 3368
Db 1501 GTCCCTCTCTGCAAAACCAATCCATGCCCTCTCTTTGTGATTCCTCGAGAAATGCTCTG 1560
Qy 3369 LeuSerGluLysGluPheTyrValAspGlnAsnValSerIleIysCysArgGluGlyPhe 3388
Db 1561 CTGTCTGAAAAGGAGTTTATTTGATGCAGATGTGTCCATCAAAATGTAGGGAAGGTTTT 1620
Qy 3389 LeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSer 3408
Db 1621 CTGCTGAGGGGACAGGCATCATTTACCTGCAACCCGACGACGACGTGGACACAGACAGC 1680
Qy 3409 AlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArg 3428
Db 1681 GCCAAATGTGAAAAAATCTCATGTGTGTCACCGAGCTCACGTAGAAAAATGCAATTTGCTCGA 1740

Qy 3429 GlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeu 3448
Db 1741 GGGTACATATTCATATATGAGACATGATCACCCTACTCATGTTACATGATGATATGTTG 1800
Qy 3449 GluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCys 3468
Db 1801 GAGGGTTTCTGAGGAGTGTGTTGTTAGAAAATGGAACATGGACATCACCTCTCTATTTCG 1860
Qy 3469 ArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCys 3488
Db 1861 AGAGCTGTCTGTCGATTTCCATGTTCAGAAATGGGGGCACTCTGCCAACGCCCAATGCTTGT 1920
Qy 3489 SerCysProGluGlyTrpMetGlyArgLeuCysGluGluProIleCysAlaIleLeuProCys 3508
Db 1921 TCTGTCCAGAGGGCTGGATGGGGCGCTCTGTGAAGAAATTAATCTGCATTTCTCCCTGT 1980
Qy 3509 LeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGlyTyrThrGly 3528
Db 1981 CTGAACGGAGTCTGCTGTGTGGCCCTTTACCAGTGTGACTGCCCGCTGGCTGGACGGGG 2040
Qy 3529 SerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArg 3548
Db 2041 TCTCGCTGTCTACAGCTGTTTGGCAGCTCTCCCTGTCTAAATGGTGGAAAAATGTGTAGA 2100
Qy 3549 ProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArg 3568
Db 2101 CCAACCGAGTCTCACTGTCTTTCTTTTGACGGGACATTAACCTGTTCAGGAAAAAGGAGG 2160
Qy 3569 ThrGlyPhe 3571
Db 2161 ACTGGGTTT 2169

Search completed: May 12, 2004, 00:15:41
Job time : 5930.35 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 10, 2004, 15:48:01 ; Search time 515.232 Seconds

(without alignments)
3846.286 Million cell updates/sec

Title: US-09-977-053-4

Perfect score: 19973

Sequence: 1 MWPLAPCCGLALVGMAT.....CHLSSWTGHCNCRKRTGF 3571

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=framet.p2n.model -DEV=xlp
-Q=/cgn2_6/ptodata/2/ina/5A COMB.seq.*
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=25
-MODE=LOCAL -OUTFMT=ptn -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977053 @CNC 1.420 @runat.06052004.075943.18213 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19958	99.9	10878	4	US-09-911-842A-1
2	16707	83.6	11230	4	US-09-911-842A-3
3	2282	11.4	1638	1	US-08-253-155A-1
4	2032	10.2	7313	4	US-09-620-312D-259
5	2027	10.1	6951	4	US-09-023-655-1265
6	2016	10.1	6951	6	5256642-1
7	2016	10.1	6951	6	5472939-1
8	1604.5	8.0	5420	6	5256642-3
9	1604.5	8.0	5420	6	5472939-3
10	1184	5.9	3934	4	US-09-023-655-1066
11	1004	5.0	8091	4	US-09-230-652-1
12	987	4.9	8257	4	US-09-484-970B-65

13	978	4.9	9723	1	US-08-083-590A-21	Sequence 21, Appl
14	978	4.9	9723	3	US-08-532-384-21	Sequence 21, Appl
15	838	4.2	763	4	US-09-484-970B-110	Sequence 110, Appl
16	834	4.2	6677	4	US-08-939-366-27	Sequence 27, Appl
17	834	4.2	6677	4	US-09-467-997-6	Sequence 6, Appl
18	763	3.8	4208	3	US-09-214-278-6	Sequence 6, Appl
19	763	3.8	4208	4	US-09-855-722-6	Sequence 3, Appl
20	762.5	3.8	3142	1	US-08-110-158-3	Sequence 3, Appl
21	762.5	3.8	3142	4	US-09-023-655-1090	Sequence 1090, Ap
22	759	3.8	5458	4	US-09-199-865-2	Sequence 2, Appl
23	759	3.8	6454	2	US-08-400-159-5	Sequence 5, Appl
24	759	3.8	6454	3	US-08-611-729A-5	Sequence 5, Appl
25	758.5	3.8	3144	5	PCT-US91-05059-1	Sequence 1, Appl
26	758	3.8	4208	4	US-09-068-740A-10	Sequence 10, Appl
27	758	3.8	5590	3	US-08-882-046-1	Sequence 7, Appl
28	742.5	3.7	4483	2	US-08-611-729A-7	Sequence 7, Appl
29	730.5	3.7	4484	3	US-08-400-159-7	Sequence 7, Appl
30	722.5	3.6	3955	3	US-09-214-278-4	Sequence 4, Appl
31	722.5	3.6	3955	4	US-09-855-722-4	Sequence 4, Appl
32	710.5	3.6	3582	2	US-08-400-159-9	Sequence 9, Appl
33	710.5	3.6	3582	3	US-08-611-729A-9	Sequence 9, Appl
34	703	3.5	376	4	US-09-833-381-205	Sequence 205, App
35	676.5	3.4	4315	3	US-08-882-046-3	Sequence 3, Appl
36	633	3.2	2508	3	US-08-981-392-1	Sequence 3, Appl
37	633	3.2	2883	3	US-08-981-392-3	Sequence 3, Appl
38	631.5	3.2	2663	4	US-09-068-740A-8	Sequence 8, Appl
39	631.5	3.2	2663	4	US-09-423-753-8	Sequence 8, Appl
40	631	3.2	2692	3	US-08-981-392-11	Sequence 11, Appl
41	628.5	3.1	5176	4	US-09-182-024A-1	Sequence 1, Appl
42	620.5	3.1	2892	1	US-08-264-534-5	Sequence 5, Appl
43	620.5	3.1	2892	1	US-08-083-590A-1	Sequence 1, Appl
44	620.5	3.1	2892	1	US-08-465-500-5	Sequence 5, Appl
45	620.5	3.1	2892	2	US-08-346-126-5	Sequence 5, Appl
46	620.5	3.1	2892	3	US-08-346-128-5	Sequence 5, Appl
47	620.5	3.1	2892	3	US-08-532-384-1	Sequence 1, Appl
48	620.5	3.1	2892	3	US-08-893-828-5	Sequence 5, Appl
49	612.5	3.1	16382	3	US-08-718-388-8	Sequence 8, Appl
50	606	3.0	5561	2	US-08-400-159-1	Sequence 1, Appl
51	606	3.0	5561	3	US-08-611-729A-1	Sequence 1, Appl
52	603	3.0	1848	4	US-09-475-460A-31	Sequence 31, Appl
53	603	3.0	1848	4	US-09-748-061A-31	Sequence 31, Appl
54	603	3.0	1878	2	US-08-435-149-17	Sequence 17, Appl
55	602	3.0	8378	5	PCT-US91-09055-1	Sequence 1, Appl
56	599.5	3.0	3339	4	US-09-423-753-4	Sequence 4, Appl
57	595	3.0	2989	6	5378464-1	Patent No. 5378464
58	592	3.0	2857	3	US-08-981-392-4	Sequence 4, Appl
59	587.5	2.9	5583	4	US-09-312-283C-372	Sequence 372, App
60	587	2.9	3231	4	US-09-866-028-14	Sequence 14, Appl
61	585	2.9	2055	3	US-08-872-855-3	Sequence 3, Appl
62	585	2.9	2800	3	US-08-872-855-1	Sequence 1, Appl
63	583	2.9	4758	3	US-09-191-647-1	Sequence 1, Appl
64	583	2.9	4758	3	US-09-540-245A-1	Sequence 1, Appl
65	583	2.9	4758	3	US-09-540-153-1	Sequence 1, Appl
66	569	2.8	373	4	US-09-621-976-11462	Sequence 11462, A
67	563.5	2.8	3834	3	US-09-209-668-18	Sequence 18, Appl
68	563.5	2.8	3854	1	US-08-365-470-1	Sequence 1, Appl
69	563.5	2.8	3858	2	US-08-344-155C-98	Sequence 98, Appl
70	563.5	2.8	3858	4	US-09-009-490A-88	Sequence 88, Appl
71	561.5	2.8	3863	4	US-08-482-073-1	Sequence 1, Appl
72	561.5	2.8	3863	6	5217870-1	Patent No. 5217870
73	558.5	2.8	1833	1	US-08-365-470-2	Sequence 2, Appl
74	549.5	2.8	2481	3	US-09-188-930-64	Sequence 64, Appl
75	549.5	2.8	2481	4	US-09-312-283C-64	Sequence 64, Appl
76	534.5	2.7	11272	4	US-09-341-461-1	Sequence 1, Appl
77	522	2.6	5499	3	US-08-479-722B-1	Sequence 1, Appl
78	522	2.6	5502	5	PCT-US95-02251-17	Sequence 17, Appl
79	499.5	2.5	1981	3	US-08-981-392-26	Sequence 26, Appl
80	474.5	2.4	11009	6	US-09-845-583A-1	Sequence 1, Appl
81	469	2.3	5089	6	517197-31	Patent No. 517197
82	467	2.3	1434	4	US-09-833-381-926	Sequence 926, App
83	457.5	2.3	1760	4	US-09-495-050A-139	Sequence 139, App
84	454	2.3	2899	3	US-08-981-392-24	Sequence 24, Appl
85	434.5	2.2	14042	3	US-08-652-877-85	Sequence 85, Appl

86 434.5 2.2 14044 3 US-08-652-877-89 Sequence 89, Appl
 87 434.5 2.2 14080 3 US-08-652-877-87 Sequence 87, Appl
 88 434.5 2.2 14083 3 US-08-476-515A-83 Sequence 83, Appl
 89 434.5 2.2 14086 3 US-08-652-877-83 Sequence 83, Appl
 90 429.5 2.2 1556 1 US-08-597-545-3 Sequence 3, Appl
 91 429.5 2.2 1556 1 US-08-457-135-3 Sequence 3, Appl
 92 427.5 2.1 1158 4 US-09-142-027A-11 Sequence 11, Appl
 93 427.5 2.1 1158 4 US-09-142-027A-13 Sequence 13, Appl
 94 427.5 2.1 1573 1 US-08-597-545-4 Sequence 4, Appl
 95 427.5 2.1 1573 1 US-08-457-135-4 Sequence 4, Appl
 96 426 2.1 1491 4 US-09-495-050A-249 Sequence 249, Appl
 97 425 2.1 3448 1 US-08-296-014A-3 Sequence 3, Appl
 98 425 2.1 3448 2 US-08-596-405-3 Sequence 3, Appl
 99 425 2.1 3448 2 US-08-877-620-3 Sequence 3, Appl
 100 425 2.1 3448 4 US-09-287-368-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-911-842A-1
 ; Sequence 1, Application US/09911842A
 ; Patent No. 6656707
 ; GENERAL INFORMATION:
 ; APPLICANT: Amgen Inc.
 ; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
 ; FILE REFERENCE: 01017/37592
 ; CURRENT APPLICATION NUMBER: US/09/911,842A
 ; PRIOR FILING DATE: 2001-07-24
 ; PRIOR APPLICATION NUMBER: US 60/222,438
 ; PRIOR FILING DATE: 2000-08-01
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 10878
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-911-842A-1

Alignment Scores:

Pred. No.: 0 Length: 10878
 Score: 19958.00 Matches: 3566
 Percent Similarity: 99.94% Conservatives: 2
 Best Local Similarity: 99.86% Mismatches: 0
 Query Match: 99.92% Indels: 0
 DB: 4 Gaps: 0

US-09-977-053-4 (1-3571) x US-09-911-842A-1 (1-10878)

QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
 DB 11 ATGTGGCTCGCTGGCCCTTTTGTGCTGGGCTCTGGCGCTCGCTTTTCGGGCTGGCGACC 70
 QY 21 PheGlnGlnMetSerProSerArganPheSerPheArgLeuPheProGluThrAlaPro 40
 DB 71 TTTCAGCAGATGTCCTCCGCGCGCAATTTTCAGCTTCGCTCTCCCGGAGACCGCGCC 130
 QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
 DB 131 GGGGCCCCGGGAGTATCCCGCGCGCGCTCTCTGGGAGCGAGCGGGGAGCAGA 190
 QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
 DB 191 GTGGAGCGGCTGGCGCAGCGTTCGCGCAGCGTGGCGCTCTCTGGGAGCTCAGCGAG 250
 QY 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAlaAsnPheArg 100
 DB 251 CCCTCGGAGCTTGTCTCTCTGGTGGATGATTTGTCAGAGGTGGCGGAAGTCAACTTCGCG 310
 QY 101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
 DB 311 ACCGAGCTCATGTTCTGTCGCGAAGCTGCTGTCCGACTTCCCGTGGTGCCCAACGCGCAG 370

QY 121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
 DB 371 CCGGTGGCCATCGTGTGACCTTCTCGTCCAAAGAACTACGTGTGTCGGCGCGTGAATTACATC 430
 QY 141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGlnIleProAlaIle 160
 DB 431 TCCACCGCGCGCGCGCCAGCAGTGGCGCTGCTCTCTCCAGAGATCCCTGCGCATC 490
 QY 161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeu 180
 DB 491 TCTACTCCGAGGTGGCGGCACCTACACCAAGGGCGCTTCAGCAAGCGCGGCAAAATTCCTT 550
 QY 181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuLeuThrAspGlyTyrSerAsn 200
 DB 551 CTTTCATGCTAGAGAAACTCAAAAGTGTGATTTCTCATCTCATGTGATATTCCTCAT 610
 QY 201 GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr 220
 DB 611 GGGGAGACCCCTAGACCAATTGCGCGTCACTGCGAGATTCAGAGTGGAGATCTTCACT 670
 QY 221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
 DB 671 TTTGGCATATGCGAGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 730
 QY 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260
 DB 731 GAGCACTGTTTACTGCTGTACACAGTTTGAAGAAATTTAGAGCTTTAGCTCCCGGCGCATG 790
 QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
 DB 791 CATGAGATCTACTTCTCGGAGTTTATTCAGATGATATGTCCTCCACTGCTCTTATCTT 850
 QY 281 CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
 DB 851 TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACA 910
 QY 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
 DB 911 GGCCATTTTGAGTGCATCTGTGAAAGGGGTATTACGGAAAGGTCTGCAGTATGAATGC 970
 QY 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyLysSerSerCys 340
 DB 971 ACAGCTTGGCCATCGGGGACATACAAACCTGAAGGCTCACCAGAGGAATCAGCAGTTGC 1030
 QY 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
 DB 1031 ATTCCATGCTCTGATGAAATCACCCTCTCCACCTGGAGGACATCCCTCGAAGACTGT 1090
 QY 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
 DB 1091 GTCTGAGAGAGGAGATACAGGGGATCTGGCCAGACTGTGAACTTGTCCACTGCGCTGCC 1150
 QY 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
 DB 1151 CTGAAGCTTCCGGAATGTTACTTATCCAAACACTTTCACCAACCACTTCAATGCA 1210
 QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
 DB 1211 GCCTGTGGGCTCGCATGTCACTGGAATTTGATCTTTGTGGGAGCAGCAGCATCTTAATGT 1270
 QY 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
 DB 1271 CTACCCATGTTTGTGTCCTGTCAGAGGCTACTTGCAGAGTAGAACAATGCTCTCAT 1330
 QY 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
 DB 1331 CTCCGCGAGCCGGAACATGCGCACATCAGCTGTCTTACAAAGGGGAATGTTATATAGACA 1390
 QY 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
 DB 1391 ACATGTTTGGTTCCTGTGATGAAGGGTACAGACTAGAGGCGAGTGAATAGCTTACTTGT 1450
 QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500

1451 CAAGGAAACGACGAGTGGGATGGCCAGAACCCGGTGTGTGGAGCGCACTGTGCCACC 1510
501 PheGlnMetProLysAspValIlelleSerProHisAsnCysGlyLysGlnProAlaLys 520
1511 TTTTCAGATGCCCAAGATATCATCATATCCCCCACAACTGTGGCAAGCAGCCAGCCAAA 1570
521 PheGlyThrIleCysPyrValSerCysArgGlnGlnGlyPheIleLeuSerGlyValLysGlu 540
1571 TTTGGACCATCTGCTATGTAGTTGCGCGCAAGGGTTCAATTTTATCTGGAGTCAAGAA 1630
541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
1631 ATGCTGAGATGTACCACTTCTGGAAATGGAATGCGAGTTCAGGCGAGCTGTGTGTAA 1690
561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
1691 GACGTGGAGGCTCTCAAAATCAACTGTCTTAAAGCATAGAGGCTAAGACTCTCGAAGCAG 1750
581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
1751 CAAGATTTCTGCCAATGTACCTGGCAGATTCACACAGCTTAAAGACAATCTGGTGAAG 1810
601 ValSerValHisValHisProAlaPheThrProProTrpLeuPheProIleGlyAspVal 620
1811 GTGTCACTCCAGTTCATCCAGCTTTCCACCCCACTTACCTTTTCCAGCTGGAGATGTT 1870
621 AlaIleValThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
1871 GCTATCGTATACACGGCAACTGACCTTATCCGGCAACGAGCCAGCTGCAATTTTCCATATC 1930
641 LysValIleAsnAlaGluProProValIleAspTrpCysArgSerProProValGln 660
1931 AAGGTATTGATGCAACCACTGTGATAGACTGGTGCAAGATCTCCACCTCCGCTCCAG 1990
661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
1991 GTCTCGGAGAGTACATGCGCAAGCTGGGATGAGCCTCAGTCTTCAGACAACCTCAGGG 2050
681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
2051 GCTGAATTGGTCTATTACCAAGAGTCTATACACAGGAGACCTTTTCCCTCAAGGGGAGACT 2110
701 IleValGlnThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
2111 ATAGTACGTATACACGCACTGACCCCTCAGGCATTAACAGGACATGTGATATCCATATT 2170
721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
2171 GTCATAAAGGTTCTCCCTGTGAAATTCACATTCACACCTGTAAATGGGANTTTTATATGC 2230
741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTrpAspPheThr 760
2231 ACTCCAGATAATACTGGAGTCAACTGTACATTAACCTGTGCTGGAGGGCTATGATTTTCA 2290
761 GluGlySerThrAspLysTrpCysAlaTrpGluAspGlyValTrpLysProThrTrp 780
2291 GAAGGCTCTACTGACAGTATTATTGTGCTTATGAAGATGGGCTCTGGAAACCAATAT 2350
781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
2351 ACCACTGAATGGCAGACTGTGCCAAAACAGTGTTCCTTAAACCAACGGGTTCAAGTCTCTT 2410
801 GluMetPheTrpLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
2411 GAGATGTTCTACAAAGCAGCTGTGTGTGATGACACAGATCTGATGAAGAAGTTCCTGAA 2470
821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
2471 GCATTTTGACAGCCCTGGGAAAATGTGTCCTCCATCATTTTGTAGTGTATGCAGAGGACATT 2530
841 AspCysArgLeuGluLeuThrLysLysTrpCysLeuGluThrLysAsnTrpAspTrp 860

2531 GACTGCAGACTGGAGGAGAACTGACCAAAAAATATTTCCTAGATAATAATTATGACTAT 2590
861 GluAsnGlyPheAlaIleGlyProGlyGlyValaIleAlaAsnArgLeuAspTrpSer 880
2591 GAAATATGGCTTTGCAATTTGGACAGGTTGGCTGGGGTGGAGCTAATAGGCTGGATCTCT 2650
881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
2651 TACGATGACTTCTCTGGACACTGTGCAAGAAACAGCCACAGCATCGCAATGCCAAGTCC 2710
901 SerArgIleLysArgSerAlaProLeuSerAspTrpLysIleLysLeuIlePheAsnIle 920
2711 TCCGGAATTAAGAAGAGTGCCTCATATCTGACTATATAAATTAAGTTAAATTTTAAACATC 2770
921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
2771 ACAGCTAGTGTGCAATTTACCCGATGAAGAATAATATACCTTTGAATGGGAATAATCAGCA 2830
941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
2831 CGACTCTTCAGACATTTGGAACATATCACAAATAAACTGAAAAGGACTCTCAACAAGAC 2890
961 ProMetTrpSerPheGlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGlu 980
2891 CCCATGTATTCTCTTCAGCTTGCATCAGAAATACTTATAGCCGACAGCAATTTCAATTAGA 2950
981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
2951 ACAAAGAAGCTTTCCCTCTTCGACAGCAGGCTCAGTCTCGAGAGGGCGTATGTGTGTC 3010
1001 AsnCysProLeuGlyThrTrpTrpAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
3011 AATTGCCCTTTGGACCTTATTATATCTGGAACATTTCACTGTGAAAGCTGCGGATC 3070
1021 GlySerTrpGlnAspGluGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTrp 1040
3071 GGTATCTTATCAAGATGAAGAAGGCGCACTTGAGTGAAGCTTTGCCCTCTCGGGATGTAC 3130
1041 ThrGluTrpIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
3131 ACGGAATATATCATTTCAAGAAACATCTCTGATTTGTAAAGCTCAGTGTAAACAGGACAC 3190
1061 TyrSerTrpSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTrpTrpLys 1080
3191 TACTCATACAGTGGACTTGAGACTGTGGAATGTGTCCACTGGGCACCTTATCAGCCAAA 3250
1081 PheGlySerArgSerCysLysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
3251 TTTGGTTCCCGGAGCTGCTCTCTGTGTCCAGAAAACACCTCAACTGTGAAAAGAGGAGCC 3310
1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
3311 GTGAACATTTCTGCAATGTGGAGTTCTTGTCCAGAGGAAAATTTCTCGGTTCTGGGTTA 3370
1121 MetProCysHisProCysProArgAspTrpTrpGlnProAsnAlaGlyLysAlaPheCys 1140
3371 ATGCCCTGTCCACCATGTCTCTGTGACTATTACCACCTAATGAGGAGAGGAGCCCTTCTGC 3430
1141 LeuAlaCysProPheTrpGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
3431 CTGGCTGTCTCTTTTATGGAACCTACCCCATTCGTGGTTCCAGATCCATCAGCAATGT 3490
1161 SerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu 1180
3491 TCAAGTTTATAGTTCAACTTCTCAGCGCAGAGGAAAGTGTGGTGGCCCCCTGCTCTCTT 3550
1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
3551 GGACATATTAAGAAGGAGCATGAATCAGCAGTCAAGTTCAGGTTTCCATGAATGCTTCTTAA 3610
1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTrpValCysLeuCysPro 1220
3611 CCTTGCCACAATAGTGAACCTGCCAGCAACTTGGGCGTGGTTATGTTGTCTCTGTCTCA 3670

Qy 1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
Db CTTGGATATACAGGCTTAAGAGTGTGAACACAGACATCGATGAGTCAGCCCATGCTTGC 3730
Qy 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyLysPheIleCysGluCysProSerGly 1260
Db CTAACAAATGGAGTTGTAAAGACCTAGTGTGGGAATTCATTTGTGAGTGGCCATCAGGT 3790
Qy 1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsn 1280
Db TACACAGGTCAAGCGTGTGAAGAAATATAATGAGTGTAGCTCCAGTCCCTTTGTTAAAT 3850
Qy 1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
Db AAAGGAATCTGTGTGATGTGGCTGGCTATCGTTGCACATGTGTGAAAGGATTTGTA 3910
Qy 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
Db GGCTTGCATTTGTGAACAGAGTCAATGAATGCCAGTCAAAACCCATGCTTAAATATGCA 3970
Qy 1321 ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340
Db GTCTGTGAAGACCAAGTTGGGGATCTTGTGCAATGCCACCTGGATTTTGGGTACC 4030
Qy 1341 ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
Db CGATGTGGAAGAACGTCGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGTACCTGT 4090
Qy 1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db AAAGACGCTGCATAGCTTCAAGATGCTGTGTGAGCTGGCTTACAGAGTCAACACTGT 4150
Qy 1381 GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db GAAATTGAACATCAATGAATGTCAGTCTTAATCCATGATAGAAATCAGGCCACCTGTGGAT 4210
Qy 1401 GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db GAAATTAATTCATACAGTTGTAAATGTGAGCGAGATTTTCAAGATTTCTGGCATATGTCATG 4270
Qy 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db GAAACAGTCTACAGGCTTTAACTCGGATTTTGAAGTTTCTGGCATATGTCATATGTCATG 4330
Qy 1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSer 1460
Db CTAGATGSCATGCTCCCATCTCTCCATGCTCTAACCTGTACCTTCTGGATGAATCCCTCT 4390
Qy 1461 AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db GACGACATGAACATATGAACACCAATCTCTATGCAGTTGATATCGGCAGCGACAAATACC 4450
Qy 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db TTGCTCCTGCTGATTAATACCGCTGGGTCTTTATGTGAATGGCGGGAAGATAACA 4510
Qy 1501 AsnCysProSerValAsnAspGlyArgTrpHisIleAlaIleThrTrpThrSerAla 1520
Db AACTGTCCCTGGTGAATGATGGCAGATGGCATCATATGCAATCACTTGGCAAGTGCC 4570
Qy 1521 AsnGlyIleThrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db AATGGCATCTGGAAGTCTATATCATGATGGAAATTAATCTGACGGTGTGTGGCTCTCTCT 4630
Qy 1541 ValGlyLeuProIleProGlyGlyValAlaLeuValLeuGlyGlnGluAspLysLys 1560
Db GTTGGTTTGGCCCATACCTGGTGGTGGTGTGTGTCTTGGGGCAGAGACAGACAAANA 4690
Qy 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp 1580
Db GGAGAGGGATTCAGCCACAGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG 4750

Qy 1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db GACTATGTCTCTTCCACAGCGGTGAAGTCACTGGGTACCTCTCCGCCAGAGAACTC 4810
Qy 1601 SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys 1620
Db AGTAAAGGAACGTTGAGCATGGCTGATTTCTTGTTCAGGAATTTGTGGGGAAGTGAAG 4870
Qy 1621 IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db ATCGATTTCAAGAGCATATTTTGTTCGATTCGCCACGCTTAGGAGGGTCACTGCCCTCAT 4930
Qy 1641 LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db CTGAGAACTGCACTCTGAAGATTTAAAGCCAGGTTTCCAAAGTCAATCTGTCTGTGTATCCA 4990
Qy 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln 1680
Db GGCTTCCAGCTGGTGGGACCCCTGTGCAGTACTGTCTGAATCAAGGACAGTGGACACAA 5050
Qy 1681 ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
Db CCACCTCTCACTGTGAACGCAATTAGCTGTGGGTGCCACCTCTCTTGGAGAAATGGCTTC 5110
Qy 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db CATTCAGCCGATGATCTTATGCTGGCAGCACAGTAACCTACACAGTGCACAAATGGCTAC 5170
Qy 1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer 1740
Db TATCTATTTGGGTGACTCAAGGATGTTCTGTACAGATTAATGGGAGCTGGAAACGGCTTCA 5230
Qy 1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db CCATCTGCTCTGATGTCGATGAGTGCAGTTCGATGGATTCAGATTGTAGTCAGCATGCTTCT 5290
Qy 1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db TGCTTGAACGTTAGATGGATCTTACATATGTTCTATGTTCCACCGTACACAGAGATGGG 5350
Qy 1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db AAAAATCTGTCAGAACCTATAAATGTAAGGCTCCAGGAAATCCGGAATAATGGCCACTCC 5410
Qy 1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db TCAGGTGAGATTTATACAGTAGTGGTCCGAAAGTCACATTTTCGTGTGAGAAAGGATACCAG 5470
Qy 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db TTGATGGGAGTAACCAAAATCACATGTTTGGAGTCTGGAGAAATGGATCATCTAATATCCA 5530
Qy 1841 TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
Db TATGTAAAGCTGTTTTCATGTTGTTAAACCGGCTATTCCAGAAATGTTGCAATGGAGGAG 5590
Qy 1861 LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla 1880
Db TTAGCATTTACTTTTGGCGCAAGTGCATATAGTGTGTAAATAAAGGATATATCTCTGGCC 5650
Qy 1881 GlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCys 1900
Db GGTGATAAAGAAATCATCTGCTTCTGTAAACAGTCTTGTGAGTCACTTCCCTCTCTGTGTGT 5710
Qy 1901 GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly 1920
Db GAACCAAGTGAAGTGTCTTAGTCCGGAATAATAAATAATGGAATAATATATTTTGTAGTGG 5770
Qy 1921 LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro 1940
Db CTTACTACCTTCTTACTGTCATCATATTCATGCGATACAGGATACAGCTTACAGGGCCCT 5830
Qy 1941 SerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeu 1960

5831 TCCATTATTGATGCAAGGCTTCTGGCAATCTGGGACAGAGCGCCACTGCTGTCACTC 5890
1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
5891 GTCTTCTGTGGAGAACACCTGCGCATCAAGATGCTGTCAATTACGGGGAATAACTTCACT 5950
1981 PheArgAsnThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr 2000
5951 TTCAGGAACACCGCTCACTTACACTTGCAGAGAGGCTATACTCTTGTGGTGTGACACC 6010
2001 IleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSer 2020
6011 ATTGAATGCTGCGCCAGCGCAAGTAGGAGTAGAGTGACACGAGTGCCTGGCTGTCTCC 6070
2021 CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly 2040
6071 TGTGATGAGCCACCATTTGTGAGCACCGCTCTCCAGAGACTGCCCATCGGCTCTTTGGA 6130
2041 AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu 2060
6131 GACATTGCAATTCTACTGCTCTGATGGTTACAGGCTAGACAAATTCGCCAGCTTCTC 6190
2061 CysAsnAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAla 2080
6191 TGCATGCCCAGGCGCAAGTGGGTACCCCCAGAGGTCAAGACATGCCCTGTGTATAGCT 6250
2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
6251 CATTTCTGTGAAAAACCTTCATCGGTTCTCTATAGCATCTTGGATCTGTGAGCAAGCA 6310
2101 LysPheAlaAlaCysSerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
6311 AAATTTGACAGCTGGCTCAGTTGTAGCTTTTAAATGATGATGAGGCTTTGACTGAAACACC 6370
2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
6371 TCAGCAAAAGATTGAATGATGAGAGGTGGCGAGTGGAAACCTTCCGCCCATGTCATCCAG 6430
2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
6431 TGCATCCCTGTGGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCA 6490
2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
6491 AACTACAGTTTGGAGCCATGGTGGCTTACAGCTGCAACAGGGGTTCATACATCAAGGG 6550
2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis 2200
6551 GAAAGAAGAGCACCTTGGGAAGCCACAGGGCAGTGGAGTAGTCTTATACCGAGTGGCCAC 6610
2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
6611 CCGGTATCTTGTGGTGAACACCATTAAGTTTGAATGGCTTCTTGGAGCATACAACTGGC 6670
2221 ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer 2240
6671 AGAATCTTTGAGAGTGAAGTAGAGTATCAGTGTAAACCGGGCTTAAAGTCAGTGGAGGT 6730
2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
6731 CCTGTATTGTCTGCCAAGCCATCGCCACTGGCAGTGAATCCCTCTCTGATGTGTGT 6790
2261 ProLeuAspCysGlyLysProProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe 2280
6791 CCTCTGACTGTGAAAACCTCCCGCATCCAGATGGCTTTCATGAAGAGAGAAACTTT 6850
2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
6851 GAAGTAGGTTCCAGGTTTCAATTTTCTGTATAGAGGTTATGAGCTTGTGTGTGACAGT 6910
2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320

6911 TCTTGACATGTTCAGAAATCTGGCAAAATGGAATGAAGTCAAAATCCAAAGTGCATGCT 6970
2321 AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThr 2340
6971 GCCAAGTCCAGAGCGCCCTCTTGGAAACACAGCTAGTATTAAAGAGGTGACCAACC 7030
2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
7031 GAGGTAGAGTGTGACATTTCTGTAAAGAGGCGATGTCTCTGCAAGGCGCCCTCTCTC 7090
2361 LeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
7091 CTGAATGCTTGGCATCCAGCAATGGAATGACTCTTCTCCCTGTTGTGAAGATTGTTCTT 7150
2381 CysThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
7151 TGTACCCACCTCCCTTAATTTCTTGTGGTGTCCCATTCCTTCTCTGCTCTTCATTTT 7210
2401 GlySerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThr 2420
7211 GGAGTACTGTCAAGTATTCTTGTGTAGTGGGTTTTTCTTAAGAGGAAATTTCTACCA 7270
2421 LeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys 2440
7271 CTCTGCCAATCTGATGGCACTGAGCTCTCCACTGCCAGAAATGTGTTCAGTAGTAATGT 7330
2441 ProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSer 2460
7331 CCCCACTGAGGAATCCCAATGGAATCATGTGTGCAAGGCTTTCCTTATCTCAGC 7390
2461 ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys 2480
7391 ACAGCTCTCTATACCTGCAAGCCAGGCTTTGAATTTGGTGGGAATACTTACCACCTTGT 7450
2481 GlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLys 2500
7451 GGAGAAATGCTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTGAGTGGCTGAAA 7510
2501 ProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal 2520
7511 CCCAAGAGATTTGAATGGCAATTTCTTTACACGAGCTTACACTATGGACAGACCGTT 7570
2521 ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu 2540
7571 ACCTACTCTTGCACCCGAGGCTTTGGCTCGAGGTCCAGTGCCTTGAAGTCTTTAGAG 7630
2541 ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln 2560
7631 ACAGGTGATTTGGGATGTAGATGCCCATCTTGCAATGCCATCCACTGTGATTTCCCA 7690
2561 ProIleGluAsnGlyPheValGluGlyValAspTyrSerTyrGlyAlaIleIleTyr 2580
7691 CCATTTGAATGTTTGTGTAGAGGTGCAAGATTACAGCTATGGTGCCATANTCATCTAC 7750
2581 SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly 2600
7751 AGTTGCTCTCCCTGGTTCAGTGGCTGGTCTATGCCATGCAGACCTGTGAAGAGTCA 7810
2601 TrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp 2620
7811 TGGTCAAGTTCCCATCCCAATGTATGCCAATAGACTGTGGCTCCCTCTCATATAGAT 7870
2621 PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMet 2640
7871 TTTGGAGACTGTAATAACTCAAGATGACAGGGATATTTTGAGCAAGAGACGACATG 7930
2641 MetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThr 2660
7931 ATGGAAGTTCCATATGTGACTCTCCCTCTTATCATTTTGGGAGCAGTGGCTGAAAA 7990
2661 TrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet 2680
7991 TGGGAAATACAAAGAGTCTCTCTGTCTACACATTCATCAAACTTTCTGTATGGTACCATG 8050

2681 ValSerTyrThrCysAsnProGlyTyrGluLeuLeuLeuGlyAsnProValLeuLeuCysGln 2700
8051 GTTTCATACACCTGTAATCAGGATATGAACCTTCGGGGAACCTGTGTGATCTGCCAG 8110
2701 GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
8111 GAAGATGGAACTTGGAAATGGCAATGCAATCTCTGATTTCAATGAATGTGACTGGCT 8170
2721 ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln 2740
8171 ACTGCTCTGAAAATGGCTTTTGGTTTTACAGAGACTAGCATGGAAAGTGTGTGCAG 8230
2741 TyrSerCysLeuProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
8231 TATAGCTGTAAACCTGGACACATCTAGTGGGCTCTGACTTAAGGCTTTTGTCTAGAGAT 8290
2761 ArgLysTrpSerGlyValSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
8291 AGAAAGTGGAGTGGTGGCTCCCACTGTGAAGCCATTTTCATGCAAAAGCCAAATCCA 8350
2781 ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu 2800
8351 GTCATGAATGGATCCATCAAGAGCACTACACATACCTGAGCAGCTTGTACTATGAG 8410
2801 CysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn 2820
8411 TGTGACCCCGGATATGTGTGTAATGGCACTGAGAGAGAACATGCCAGGATGACAAAAAC 8470
2821 TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla 2840
8471 TGGAGTGAAGATGAGCCCATTTGCATTTCTGTGAGCTGCAGTTCACTCCCGCAGTCTGACC 8530
2841 AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn 2860
8531 AATGGCCAGTGGAGAGACGAGTACATATCCAAAAAGAGATTGAATACACTTGCAT 8590
2861 GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer 2880
8591 GAAGGGTTCTTGTCTGAGGGAGCAGAGTGGGGTTTGTCTTGGCCAAATGGAAAGTTGGAGT 8650
2881 GlyAlaThrProAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGly 2900
8651 GGAGCCATCTCCGACTGTGTGCTGTGAGTGCACCCCGCCACACTGGCCAAATGG 8710
2901 ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly 2920
8711 GTGACGGAAGGCTGGACTATGGCTTCATGAAGGAAGTAACATTCCACTGTCTATGAGGC 8770
2921 TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu 2940
8771 TACATCTTGACGGTCTCCAAACTCACCTGTCACTCAGTGGCAACTGGGATGCAGAG 8830
2941 IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
8831 ATTCTCTCTGTAAACAGTCAACTGTGACCTCTCTGAGATCTTGCCTCAATGGTTCCCT 8890
2961 AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980
8891 AATGGTTTTCTTTATTCATGGGGGCCATATACATATCAGTGTCTTCTCTGTATTAAG 8950
2981 LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerPro 3000
8951 CTCCTGGAATTCATCAAGAAAGTGCCTCTCCAAATGGCTCTCTGGAGTGGCAGCTCACCT 9010
3001 SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr 3020
9011 TCTGCTGCTGTGAGATGTTCCACACCAAGTAATTAATGAATATGTAATGTCAATGGACA 9070
3021 AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
9071 GATTTTGTGACTGTGGAAGGAGCCCGGATTCAGTGTCTCAAGGGCTTCAAGCTCTTAGGA 9130

3041 LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerClyPheProHisCysGlu 3060
9131 CTTTCTGAAATCACCTGTGAAGCCGATGGCAGTGGAGTCTGGGTTCCTCCCACTGTGAA 9190
3061 HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer 3080
9191 CACACTTCTTGTGGTTCTCTTCCAAATGATACCAATGCGTTTCATCAGTGGAGCAGCTCT 9250
3081 TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
9251 TGGAAAGGAAAAATGTGATATCTTACAGCTGCAGGCTCTGGATATGTCATACAGGCAAGTTCA 9310
3101 AspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeu 3120
9311 GATCTGATTGTACAGAGAAAGGGTATGGAGCCAGCTTATCCAGTCTGTGTAGCCCTTG 9370
3121 SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr 3140
9371 TCTGTGGGTCCCACTGCTGTGCGCAATGCGAGTGGCAACTGGAGAGGCACACCTAT 9430
3141 GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThr 3160
9431 GAAAGTGAAGTGAACCTCAGATGTCTGGAAGGTTATACGATGATACAGATACAGATACA 9490
3161 PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys 3180
9491 TTCACCTGTCAGAAAGATGGTCTGCTGGTTCCCTGAGAGAAATCTCTGCGAGTCTTAAAAA 9550
3181 CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
9551 TGTCTCTCTCCGGAAAAACATACACATATATCTTGTTCATGGGACGATTTTCAGTGTGAAT 9610
3201 ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
9611 AGCAAGTTCTCTGTCTCATGTGCAGAGGATATACCTTTTGGAGGAGTTAAATATCAGTA 9670
3221 CysGlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProValSer 3240
9671 TGTCACTGTGATGAGAACCTGGGAGCCACCATCTCCGATGATCTTGGAGTCCAGTTTCT 9730
3241 CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
9731 TGTGGAAACCTGAAAGTCCAGAACATGATGATTTGTGTTGGCAGTAAATACACCTTGAA 9790
3261 SerThrIleIleThrGlnCysGluProGlyTyrGluLeuGluCysAsnArgGluArgVal 3280
9791 AGCACAATTAATTTATCAGTGTGAGCTGTGCTATGAACTAGAGGGGAACAGGGAACGGCTC 9850
3281 CysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGlu 3300
9851 TGGCAGAGAACAGACAGTGGAGTGGAGGGGTGGCAATATGCAAGAGACAGGTGTGAA 9910
3301 ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsn 3320
9911 ACTCCACTTGAATTTCTCAATGGAAAGCTGACATTTGAAACAGGACGACTGGACCCAAC 9970
3321 ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
9971 GTGGTATATTTCTCTGCAACAGAGCTACAGTCTTGAAGGGCCATCTGAGGCACTGTCACA 10030
3341 GluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValPro 3360
10031 GAAATGGAACTCTGGAGCCACCCAGTCTCTCTCTGCAACCAATCAATGCTCTGTCTCT 10090
3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
10091 TTTGTGATTTCCCGAGAAATGCTCTGCTGTCTGAAAGAGGATTTTATGTTGATCAGATGTG 10150
3381 SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnPro 3400
10151 TCCATCAATATGTAAGGAAGGTTTCTGCTGTCAGGGCCACGGCATCATTTACTCTGCAACCC 10210
3401 AspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAla 3420

DB 10211 GACGAGCGTGGACACACAGCGCCAAATGTGAAAATCTCATGTGTCACACAGCT 10270
QY
QY 3421 HiaValGluAenAlaAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
DB 10271 CACGTAGAAAATGCAATTTGCTCGAGGCGTACATTTATCAATATGAGACATGATCACCTAC 10330
QY 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAenGly 3460
DB 10331 TCATGTTACAGTGGATACATGTTGGAGGGTTTCTGAGGAGTGTGTTGTTAGAAAATGGA 10390
QY 3461 ThrTyrThrSerProPheLeuCysArgAlaValCysArgPheProCysGlnAenGlyGly 3480
DB 10391 ACATGACATCACTCTTATTTGACAGAGTGTCTGTGATTTCCATGTCAGAAATGGGGC 10450
QY 3481 IleCysGlnArgProAenAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGlu 3500
DB 10451 ATCTGCCAACGCCCAATGCTTGTCTGTCTCAGAGGGCTGGATGGGGCGCTCTGTGAA 10510
QY 3501 GluProPheCysIleLeuProCysLeuAenGlyGlyArgCysValAlaProTyrGlnCys 3520
DB 10511 GAACCAATCTGCATTTCTTCTGTCTGACAGAGTGTCTGTGAGGGCTTACAGTGT 10570
QY 3521 AspCysProGlyTyrThrThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
DB 10571 GACTGCCCGCTGGCTGACGGGCTCTCGCTGTCTACAGCTGTTTGCAGTCTCCCTGC 10630
QY 3541 LeuAenGlyGlyCysValArgProAenArgCysHisCysLeuSerSerThrThrGly 3560
DB 10631 TTAATAGTGTGAAAATGTTGTAAGACCAACCGATGTCATGCTCTTCTTCTTGGACGGGA 10690
QY 3561 HiaAenCysSerArgGlyArgArgThrGlyPhe 3571
DB 10691 CATAACTGTTCCAGGAAAAGAGGACTGGGTTT 10723

RESULT 2

US-09-911-842A-3
; Sequence 3, Application US/09911842A
; Patent No. 6658707
; GENERAL INFORMATION:
; APPLICANT: Angen Inc.
; TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF
; FILE REFERENCES: 01017/37592
; CURRENT APPLICATION NUMBER: US/09/911.842A
; CURRENT FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: US 60/222,438
; PRIOR FILING DATE: 2000-08-01
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 11230
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-911-842A-3

Alignment Scores:

Pred. No.: 0 Length: 11230
Score: 16707.00 Matches: 2915
Percent Similarity: 89.72% Conservative: 289
Best Local Similarity: 81.63% Mismatches: 361
Query Match: 83.65% Indels: 6
DB: 4 Gaps: 5

US-09-977-053-4 (1-3571) x US-09-911-842A-3 (1-11230)

QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTyrAlaThr 20
DB 176 ATCTGTTGCGCGCTGGCTTTTGTGCTGCGCTCTGCGCACTGGGTGTGGGCTGGACCAAC 235
QY 21 PheGlnGlnMetSerProSerArgAenPheSerPheArgLeuPheProGluThrAlaPro 40
DB 236 TTCACGCCCTGGCGCCCTCGCTCAACTTCAGCTTCGCGCTGTTCGCCGAGGCTCTCTCG 295

QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
DB 296 GGGGCTCTGGCAGACTGGCGGTACTCTCCGCGTCCAGTGAGGAGGAGGAGGAGGAGG 355
QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSer 79
DB 356 AAAGTGAGGAGCGCTGGCGCGCGTTCGCGAGCGCGTTCGCGCGGAGCTCAGC 415
QY 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerSerValGlyGluValAenPhe 99
DB 416 GCGAGCTTGGAGCTGCTCTCTGTTGAGCAGCTGCTCCAGCGTGGCCCAACCAACTTC 475
QY 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAla 119
DB 476 CTCACGAGCTCAAGTTGTTGGCAGCTGTTCCGACTTCCCGTGGTGTCCAGGCG 535
QY 120 ThrArgValAlaIleValThrPheSerSerIysAenTyrValValProArgValAspTyr 139
DB 536 ACGGTGTGGCCATCGTCACTTCTCATCAAGAACCAACGTGTGTGGCGCGCGTGGATTAC 595
QY 140 IleSerThrArgArgAlaArgGlnHisCysCysAlaLeuLeuGlnGluLeuProAla 159
DB 596 ATCTCCACGCGCGCGCACCAACCAAGTGGCGCTGCTCAGCGCGAGATCCCGGCG 655
QY 160 IleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
DB 656 ATCACTACCGGTGTGGCAGCTTATACCAAGGGCGCTTCCAGCAAGCGCGCAATC 715
QY 180 LeuLeuHisAlaArgGluAenSerThrLysValValPheLeuIleThrAspGlyTyrSer 199
DB 716 CTTCGTCACTCTAGAGAAAACCTCCACCAAGTCAATTTCTCATCCAGCGCGCTATTTC 775
QY 200 AenGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
DB 776 AATGGCGAGACCCAGACCTATTTCAGCATTCGCTTCGGGATTTCCGAGTGGAGATCTTC 835
QY 220 ThrPheGlyIleTrpGlnGlyAenIleArgGluLeuAenAspMetAlaSerThrProLys 239
DB 836 ACGTTCCGGATTTGGCAGGGGAAATATCCGGGAACCTGAATGACATGGCTTCCACCCGAG 895
QY 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
DB 896 GAAGACATTTGTACCTGCTCCACAGTTTTGAAGAAATTTAGGCTTTAGCTCGCAGGCG 955
QY 260 LeuHisGluAenLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
DB 956 TTGCATGAAGATCTACCTTCTGGGAGTTTATCAAGAGGATATGGCCACATGCTCTTAT 1015
QY 280 LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis 299
DB 1016 CTCTGTGAGGCTGGGAAAAGACTGCTGTGACAGAAATGGCCAGCTGCAAAATGTGGACAC 1075
QY 300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGlu 319
DB 1076 ACGGTCATTTGAATGATCTGTGGAAGAGGCTTATTACGGGAAAGGTTCTCGAGCATGAG 1135
QY 320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
DB 1136 TGCACACTTCCCATCAGGACATATAGCCGAGCTTCTCCAGGAGGATCAGACAC 1195
QY 340 CysIleProCysProAspGluAenHisThrSerProGlySerThrSerProGluAsp 359
DB 1196 TGCATCCATGCTCTGACGTAAGGCACACCTCCCGACCTGGAAGCCTTCCCTCGAAGAC 1255
QY 360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
DB 1256 TGGGTGTGGCGAGAGGATACAGAGATCTCGCCAGACCTGTGAGGTTGTCCACTGTCTCT 1315
QY 380 AlaLeuLysProProGluAenGlyTyrPheIleGlnAenThrCysAenHisPheAen 399
DB 1316 GCCCTGAAGCCTCTGAAATGGTTTTTATACAAACACTTGCAGAAAACCTACTTCAAT 1375
QY 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeu 419

Db 5750 GGGTCGGCATTCACATATGCGAGTAAGGTGGTGTACAGGTGTGTATTAAGGATATACCTTTG 5809
QY 1880 AlaGlyAspLysGluSerCysLeuAlaAsnSerSerTyrTrpSerHisSerProProVal 1899
Db 5810 TCTGGGGATGAAGAGTCAGATCCCTTGTCTGTAGTGGTTCCTGGAGTCATCTCTCTCTGTG 5869
QY 1900 CysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSer 1919
Db 5870 TCGGGCTAGTGAAGTGTTCGCCAGCTGAGGACATAAATAAGCGCAATACATCTTAAGT 5929
QY 1920 GlyLeuThrTyrIleuSerThrAlaSerTyrSerCysAspThrGlyTyrSerIleuGlnGly 1939
Db 5930 GGGCTCACCTATCTTATATGATCGTACTCTCTGTGAGAACGATACAGTTTACAGGGC 5989
QY 1940 ProSerIleIleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHis 1959
Db 5990 CCATCCCTCTTGAAATGCACAGCTTCGGCAGCTGGGACGAGCGCCACCTAGCTGTCAA 6049
QY 1960 LeuValPheCysGlyGluProProAlaIleLeuAspAlaValIleThrGlyAsnAsnPhe 1979
Db 6050 CTTGTCTCTCGGAGAGCTCCCAATCGTCAAGATGTGTCTCATCTCGGAGCAACTTC 6109
QY 1980 ThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrIleuAlaGlyLeuAsp 1999
Db 6110 ACTTTTGGGAACACAGTTGCTTACATGCAATGAGGGCTTACACCTTGTCTGGGCTGAC 6169
QY 2000 ThrIleGluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaVal 2019
Db 6170 ACCATCATATGCCGAGCCCAAGCGCAATGGAATTCAGTACCAACCAAGTGCCTGGCTGTC 6229
QY 2020 SerCysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhe 2039
Db 6230 TCTGTGACGAGCGCCGCCAATGTGACACCGCTCTCCAGAGACTGCTCACAGGCTCTTT 6289
QY 2040 GlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeu 2059
Db 6290 GGAGACACCGCGTTTACTACTGTGGGATGGCTACAGCGCTGGCTGATTAATTCGCCAGTC 6349
QY 2060 LeuCysAsnAlaGlnGlyLysTyrValProProGluGlnAspMetProArgCysIle 2079
Db 6350 ATCTGCANATGCCAGGGAACTGGGTTCGCCCGCGGCGAGCTGTGCGCGCTGCATA 6409
QY 2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLys 2099
Db 6410 GCTCACTCTGTGAAAAACCCCATCTGTTCTACAGCATCTTGGAACTGTGAGCAAA 6469
QY 2100 AlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsn 2119
Db 6470 GCAAAAGTTTGCAGCTGGCTGGTAGTGAGCTTCAGTGATGAGGGTTTGTGCTGAAC 6529
QY 2120 ThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIle 2139
Db 6530 ACCTCAGCGAAGATTGATGCTGAGAGGTGGAGGTGGAGCCCTCTCCCTCTGGTC 6589
QY 2140 GlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGly 2159
Db 6590 CAGTGATCCCGTGGCGATGCGAGAGCTCCCAAGCATCGCAATGGCTACCCGAGTGGG 6649
QY 2160 SerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLys 2179
Db 6650 ACAAACTACAGTTTGGGCGCGTGGTGGCTTACAGCTGCCAACAGGGATTCATATCAAG 6709
QY 2180 GlyGluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCys 2199
Db 6710 GGGGAGAGAGAGACAGTGTGAGGCCACAGGACAGTGGAGTAAACCCACGCCCACTGTC 6769
QY 2200 HisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThr 2219
Db 6770 CATCTGTGCTGTAAACAGGCGACCTAAGGTTGAGAACGGCTTCTCTGGAGCACACCACT 6829
QY 2220 GlyArgIlePheGluSerGluValAlaArgTyrGlnCysAsnProGlyTyrLysSerValGly 2239
|||||

Db 6830 GGCAAGACCTTTGAGAGCGCAAGCAAGGTTCCAGTGCAAACCCAGGCTTATAAGGACGCCGA 6889
QY 2240 SerProValPheValCysGlnAlaAsnArgHisTyrPheHisSerGluSerProLeuMetCys 2259
Db 6890 AGTCTGTGTGTTGTTGGCAGCAATGCACTGGCACAGCGAGCCCTCTGTCTGCTGC 6949
QY 2260 ValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsn 2279
Db 6950 ACCCTCTCAACTGTGGAAACCCCTCCCATTCAGAAATGGCTTTTGAAGAGGAAGC 7009
QY 2280 PheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAsp 2299
Db 7010 TTTGAAGTAGGGTCCAGGTTTCAGTTTCTCTAATGAGGATATGAGCTCTGTGGTAT 7069
QY 2300 SerSerTyrThrCysGlnLysSerGlyLysTyrPheAsnLysLysSerAsnProLysCysMet 2319
Db 7070 AATTTCTGGACTTGGCAAAATCTGGCAATGAGTAAGAACCCAGCCCGAAGTGTGTC 7129
QY 2320 ProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThr 2339
Db 7130 CCACCAAGTGTGCAGAGCTCTCTCTTTAGAAACAGCTCGTATTGAAGGAATTAGCT 7189
QY 2340 ThrGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSer 2359
Db 7190 TCCAGGTAGAGTGTATGACCATTTCTGTAAAGGGGATGCTCTTGCAGAGGCCCTCT 7249
QY 2360 ValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleVal 2379
Db 7250 GTCTGAAGTGTCTGCATCGCGCAATGGAATGGTTCTTTCTTATTTGAAGATGTC 7309
QY 2380 LeuCysThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHis 2399
Db 7310 CTTTGTCCCTCGGCTCTCTTGAATTCCTTCGGGCTCTCTCGCTCTTCGGTGTCTTCAT 7369
QY 2400 PheGlySerThrValLysTyrSerCysValGlyPhePheLeuArgGlyAsnSerThr 2419
Db 7370 TTTGGCAGTACTGTCAAGTATCTGTGTGTCAGCGGTTTTCTTAAGAGGCGATCCAAAC 7429
QY 2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
Db 7430 ATCTCTGCGAGCTGATAGCCTCGAGTGTCTCATTTGCCGAATGCGTTCCGGTAGAA 7489
QY 2440 CysProGluProGluGlnIleProAsnGlyIleLeuAspValGlnGlyLeuAlaTyrLeu 2459
Db 7490 TGTCCCAACCTGAGGAGATCTCAACGGTATCATCCACGTACAGGGCTTGCCTATCTC 7549
QY 2460 SerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeu 2479
Db 7550 AGCACCACTGTACACTGCAGCCAGCGCTTTGAGTTAGTGGCAATGCTTACCACCTC 7609
QY 2480 CysGlyGluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeu 2499
Db 7610 TGTGGGAAAAATGCCAGTGGCTCGGAGGAAAAACCAATGTGCAAAACCATTTGAATGCCCA 7669
QY 2500 LysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr 2519
Db 7670 GAGCCCAAGAGATTTTAAATGGCCAAATCTCTTCCGTGAGCTTTCAGTATGAGCAAAAC 7729
QY 2520 ValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
Db 7730 ATCATATCTTTTGTGACCGGGCTTCGGGCTCGAGGTCCCAATCCCTGACCTGTTTA 7789
QY 2540 GluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerPro 2559
Db 7790 GAGACAGGTGACTGGGATATGATGCCCTCTTGTGATGCCAATCCATGCGAGTGAGCCCA 7849
QY 2560 GlnProIleGluAsnGlyPheValGluGluValAspTyrSerTyrGlyAlaIleIle 2579
Db 7850 CAGCCCATTTGAAATGGTTTCGTAGAGGTGGGATTTACAGATACGGTGCATGATCATC 7909
QY 2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
Db 7910 TATAGTGTCTCTCGGTTTCAGGTGCTTGGTTCATGCCATGCCAGCTGTGAAGAGTGC 7969

2600 GlyTyrSerSerIleP-orthrCysMetProIleAspCysGlyLeuProProHisIle 2619
Db GATGTGTCAGCTCCAGCCCAACCTGTGTACCCATAGACTGCGGTCTCCCTCCTCACATA 8029
2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAsp 2639
Db GACTTGGTGACTGTACTAAAGTCAGAGATGGCCAGGACATTTGATCAAGAGATGAC 8089
2640 MetMetGluValProTyrValThrProHisProTyrHisLeuGlyAlaValAlaLys 2659
Db ATGATGGAAGTCCCATCTGGCT-----CACCCCTCAACATTTGGAAGCAACAGCTAAG 8143
2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhelLeuTyrGlyThr 2679
Db GCCTTGGAAATAACAAGAGATGCGCTGCCTCATATGCATCCCATCTCTCTATGGCAGC 8203
2680 MetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLysCys 2699
Db ATGTTTCTACAGCTGCGAGCTGTGTATGAACCTGCTGGGAATCCCTGTGTGCTGATCTGC 8263
2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
Db CAGGAAGATGGTACGTGAATGGTACCGCACCTCTTGCATTTCCATTTCAATGATGATTG 8323
2720 ProThrAlaProGluAsnGlyPheLeuA-9PheThrGluThrSerMetGlySerAlaVal 2739
Db CTTGTGTCTCCCGAAATGGCTTTTACATTTTACACACAGACGACTATGGGCACTGTGCA 8383
2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
Db CAATATAGCTGCAAGCGGGGCACATTTAGAGGCTCCCACTTAAGACTCTGTCTGTCAG 8443
2760 AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
Db AATAAGCAGTGGAGTGGCACTGTGTCCACGCTGTGAAGCCATCTCATGCGATGAAGCCAAAC 8503
2780 ProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyr 2799
Db CCATCTGGAATGGATCCATCAAGAGAGATGACTACTCTTACCTGGGTGTGTATACTAC 8563
2800 GluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLys 2819
Db GAGTGTGACTCTGGCTATATTCTCAATGGCTCTAAGAAGAGGACATGCCCAAGAAATAGA 8623
2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
Db GATTGGATGGGCATGAGCCCATGTGTATTCTGTAGACTGTGGCTCACCCCCAGTCCCC 8683
2840 AlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
Db ACCAATGGCCGAGTAGGAGGAGAGAGATACATTCCAAGAGAGATACATCTCTTGC 8743
2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
Db CTTGAAGGGTTTCATCTGGAAGGAGCCAGAGTGTGTATCTGTCTTACCAATGGAAGTTGG 8803
2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProGluGlnLeuAlaAsn 2899
Db AGTGTGGCCACTCCCAAGCTGCATGCTGTATAGATGTCTTGTAGATGTCTTCCACTGTCTAGAG 8923
2900 GlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu 2919
Db GGGGTGGGAGTGGCTTAGCATATGGGTTCAAGAGAGAGATAGCGTTCCACTGTCTAGAG 8923
2920 GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
Db GCTATGTCTCAGGGGGCTCCAAAGACTCACCTGTCACTGCCAAATGGGACTTGGGATGCA 8983
2940 GluIleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhe 2959
Db GAAGTCCCTGTCTGTAAACAGTACTCTGTGTGTCTCTCCGCGACCTCTCCCGAGGCTTC 9043

2960 ProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyr 2979
Db CTTAATGGCTTTTCTTTTATCATGGGGCCACATACATGATATAGTGTTTTACTGTTAT 9103
2980 LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
Db AAGCTTCATGGAACCCATCAGAGAGATGCTTCCCAATGGTCTCTGGAGCGCAGCTCG 9163
3000 ProSerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGly 3019
Db CCATCTGCTTACCTTTCAGGTGTTCACACCCATCATCAACAGGGAACCATCAACGCA 9223
3020 ThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeu 3039
Db ACTGATTTGGAGTGTGGAAGACGGTCCAGATGAGTGTCTTCAAGGCTTCAAGCTGTGT 9283
3040 GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPhePheProHisCys 3059
Db GGAATTTCTGAATCACCCTGTGATGATGCAATGGCCAAATGGTCTGAC---GTCCCACTGTGT 9340
3060 GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer 3079
Db GAGCAGCTCAGTGGCGGCTCTCCCAACCATATCCCAACGCAATTTGCTTTGAGGGCAGC 9400
3080 SerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySer 3099
Db CTTTGGAGGACAAATGTGTAACTTACAGTTCAGACTGAGCTTACCATCATGCAAGTAGT 9460
3100 SerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluPro 3119
Db TCAGATCTGATTTGTACGGAAAAAGGATATGAGCAGCCAGCTTACCACAACTGTGAACCC 9520
3120 LeuSerCysGlySerProSerValAlaAsnAlaValalathrGlyGluAlaHisThr 3139
Db CTGTCTGTGGACCCCACTGAGTGTAGCAATGTCAGTGGCAACAGAGAGGCTCATACC 9580
3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp 3159
Db TATGAAGCAAGTGAATCAGTGTCTGGAAGGATATGTGATGATTCGGATACAGAT 9640
3160 ThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLys 3179
Db ACATTCACCTGCCAGCAAGATGGCCATTTGGTCCCTCGAAAGAATCACTGCACTCTAAA 9700
3180 LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerVal 3199
Db AATGCCCTGTGCCATCCCAACATGACGATGCTTTTCCAGGAGATGACTTCCAGGTG 9760
3200 AsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSer 3219
Db AACAGACAGTTTCTGTGTCACTGTGCAAGAGGTTTACCCACAGAGAGTGAAGTGTCA 9820
3220 ValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProVal 3239
Db ACATGCCAGCCCGCAGCTATCGGAGCCACCATTTTCTGATGAATCTCTGATCCAGTT 9880
3240 SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPhe 3259
Db GTTTGTGGGCATCTTCAAGAGCCAGCGCATGCTGCTGGTGTGGCAATAGCACAGTTT 9940
3260 GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg 3279
Db GGAAGCACCATTTGTTTACCAGTGTGACCTTGGCTACAAATAGAGGGGACAGGACGA 10000
3280 ValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLysGluThrArgCys 3299
Db ATCTGCCAGGAGAACAGACAGTGGAGAGAGTGGCAGTGTGCAGAGAACAGATGT 10060
3300 GluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyPro 3319
Db GAGACTCCAGTGTGATTTCCCAATGGAGAGGCTGTCTGGAAACACCAATCTTGGACCC 10120
3320 AsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCys 3339

```
Db 10121 AGCTCTCTGTTTCTGTCACAGAGGCTACACCTCGGAGGGTCCCGGAGGCACTGTC 10180
Qy ThrGluAenGlyThrTrpSerHisProValProLeuCysLeuProAsnProCysProVal 3359
Db 10181 ACTGCAATGGACCTGATCCCTGACTCCCTCTGCAACCAATCCATGCTGTC 10240
Qy ProPheValIleProGluAenAlaLeuLeuSerGluLeuGluPheTyrValAspGlnAen 3379
Db 10241 CCTTTTGTGATCTCTGAGAACGCGCTCTCTGAAAAAGAGTTTTATGTGCACAGAA 10300
Qy ValSerIleLeuCysArgGluCysPheLeuGluGlnGlyHisGlyIleLeuThrCysAen 3399
Db 10301 GTATCTATCAATGTCAGGAGGAGCTCTCTGCTCAAGGCAATGGTGTATCATCGTGCAGC 10360
Qy ProAspGluThrTrpThrGlnThrSerAlaLeuCysGluLeuIleSerCysGlyProPro 3419
Db 10361 CTGACGAGACATGACGACACCAATGCCAGATGTGAAAAATCTCTGTGTGCTCTCCA 10420
Qy AlaHisValGluAenAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThr 3439
Db 10421 AGTCACGTAGAAATGCAATGCTCTGAGGAGTGTATTACCAATGAGGACATGATCACC 10480
Qy TyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAen 3459
Db 10481 TACTCTCTGTACAGTGGCTACATGCTAGAAAGTGTCCCTCCGAGTGTTCCTCTAGAAAAT 10540
Qy GlyThrTrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAenGly 3479
Db 10541 GGAACATGGACACCATCTCTGTTGACAGAGTGTCTGTGCTGCTCCATGTCAGAAATGGA 10600
Qy GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCys 3499
Db 10601 GGTGTCTGTCAACGCTCAATGCTGTCTCATGCCAGAGGCTGGATGGAGAGTCTCTGT 10660
Qy GluGluProIleCysIleLeuProCysLeuAenGlyGlyArgCysValAlaProTyrGln 3519
Db 10661 GAAGAGCAATATGATCATCTCCCTGTTTGAATGGTGGGCGCTGTGTGCCCCCTTATCAG 10720
Qy CysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerPro 3539
Db 10721 TGTGACTGCCCCACAGGCTGGAGTGGTCCCGCTGTCTATACAGTACTTGTCTCAGTCCCG 10780
Qy CysLeuAenGlyGlyCysValArgProAsnArgCysHisCysLeuSerSerTrpThr 3559
Db 10781 TGTCTAAATGGGGGAAATGCAATAGACCAACCGATGCTATGTCTCTCAGCCTGGACA 10840
Qy GlyHisAenCysSerArgIleArgThrGly 3570
Db 10841 GGACATGATGTTTCCAGGAAAGGAGCGCGG 10873
```

RESULT 3

US-08-253-155A-1

; Sequence 1, Application US/08253155A

; Patent No. 5691147

; GENERAL INFORMATION:

; APPLICANT: Gyuris, Jeno

; APPLICANT: Draetta, Giulio

; TITLE OF INVENTION: CDK4 Binding Proteins

; NUMBER OF SEQUENCES: 95

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/253,155A
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1638 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-253-155A-1
```

Alignment Scores:

```
Pred. No.: 1.88e-179 Length: 1638
Score: 2282.00 Matches: 413
Percent Similarity: 97.65% Conservative: 3
Best Local Similarity: 96.95% Mismatches: 2
Query Match: 11.43% Indels: 8
DB: 1 Gaps: 0
```

US-09-977-053-4 (1-3571) x US-08-253-155A-1 (1-1638)

```
Qy 3154 MetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArg 3173
Db 16 ATGGATACAGATACAGATACATTCCTCTCCCGAAAAATAACACATATATCTGTGAGAGA 75
Qy 3174 IleSerCysSerProLysLysCysProLeuProGluAenIleThrHisIleLeuValHis 3193
Db 76 ATCTCTGCGAGTCTAAAAAATGCTCTCTCCCGAAAAATAACACATATATCTGTGAGAGA 135
Qy 3194 GlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPhe 3213
Db 136 GGGGAGCATTTTCAGTGTGTAATAGGCAAGTTTCTGTGTCTATGTGCAAGAGGATATACCTTT 195
Qy 3214 GluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAsp 3233
Db 196 GAGGAGGATTAACATATCAGTATGTGAGTGTGAGACCTGGGAGCCACCATTTCTCCGAT 255
Qy 3234 GluSerCysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValVal 3253
Db 256 GAATCTTGCAGTCCAGTTTCTTGTGGGAAACCCAGAAAGTCCAGAACATCGATTTGTGGTT 315
Qy 3254 GlySerLysTyrThrPhe-GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLe 3273
Db 316 GGCAGTAAATACACCTTTTGAAGACAAATTTATTCAGTGTGAGCTTGGCTATGAACT 375
Qy 3273 uGluGlyAsnArg-GluArgValCysGlnGluAenArgGlnTrpSerGlyGlyValAlaI 3293
Db 376 GAGGGGGAACAGCGCAACCTGTCTCCAGAGACAGACAGTGGAGTGGAGGGGTGGCAA 435
Qy 3293 leCysLysGluThrArgCysGluThrProLeuGluPheLeuAenGlyLysAlaAspIleG 3313
Db 436 TATGCAAAAGAGACAGGAGTGTGAAATCTCCATTTTCAATGGGAAAGCTGACATTG 495
Qy 3313 LuAenArgThrThrGlyProAenValTyrSerCysAenArgGlyTyrSerLeuGluG 3333
Db 496 AAAACAGGACGAGCTGGACCCCAAGTGTATATTTCTTCCAAACAGAGGCTACAGTCTTGAAG 555
Qy 3333 lYProSerGluAlaHisCysThrGluAenGlyThrTrpSerHisProValProLeuCysL 3353
Db 556 GGCCATCTGAGGACACATGTCACAGAAAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGCA 615
Qy 3353 yProAenProCysProValProPhe-ValIleProGluAenAlaLeuLeuSer-GluTy 3372
Db 616 AACCAAAATCCATGCTGTTCTTTTGGTGAATTTCCCGAGATGCTCTGTCTGTCTTGAAGA 675
```

3372 sGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnG1 3392
Db |||||
676 CGAGTTTATGTTGATCAGATGTGTCATCAATGTAGGAGGTTTCTGTCGAGGG 735
3392 yHisGlyIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysG1 3412
Db |||||
736 CCACGGCATCATTACCTGCAACCCGACGAGAGCTGGACACACAGCAAGCCCAATGTGA 795
3412 uLysIleSerCysGlyProProAlaHisVal-GluAsnAlaIleAlaArgGlyValHisT 3432
Db |||||
796 AAAAATCTATGTGGTCCACCACTCAGGTAGCAAAATGCAATGTCTGAGGGGTACAT 855
3432 yGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheL 3452
Db |||||
856 ATCAATATGAGACATGATCACTTCTCATGTTACAGTGGATACATGTTGGAGGTTTC 915
3452 euArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValC 3472
Db |||||
916 TCAGGAGTGTGTTGTAGAAAATGGAACATGGACATCACCTCTTAATTTGCAGAGCTGTCT 975
3472 yArgPheProCysGln-HenGlyGlyIleCysGlnArgProAsnAlaCysSerCysPro 3491
Db |||||
976 GTGCAATTCATGTCAAGAAATGGGGGATCTGCCAAGCCCAAAATGCTTCTCTGTCCA 1035
3492 GluGlyTrpMetGlyArgLeu-CysGluGluProIleCysIleLeuProCysLeuAsnG1 3511
Db |||||
1036 GAGGGCTGATGGGGGCGCTCTGTGAGAGAACCAATCTGCAATCTTCCCTGTCGAAACG 1095
3511 yGlyArgCysValAlaProTyrGlnCysAspCysProGlyTrpThrGlySerArgC 3531
Db |||||
1096 AGGTGCTGTGTGGCCCTTACCAGTGTGACTGCCCGCTGGCTGGACGGGGTCTGGCTG 1155
3531 sHisThr-AlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnA 3551
Db |||||
1156 TCAACAACTGTGTGCCAGTCTCCCTGTCTTAATGTGGAAATGTGTAAGACCAAC 1215
3551 rGlyHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyP 3571
Db |||||
1216 GATGTCATGTCTTCTTCTTGACGGGACATTAATCTTCCAGGAAAGGAGGACTGGGT 1275
3571 he 3571
1276 TT 1277

RESULT 4

US-09-620-312D-259
; Sequence 259, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aizong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungling
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 259
; LENGTH: 7313
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (13)..(6147)
; US-09-620-312D-259

Alignment Scores: Length: 7313
Pred No.: 159e-157 Matches: 564
Score: 2032.00 Conserv: 265
Percent Similarity: 40.13% Mismatches: 831
Best Local Similarity: 27.30% Indels: 406
Query Match: 10.17% Gaps: 85
DB: 4

US-09-977-053-4 (1-3571) x US-09-620-312D-259 (1-7313)

QY 1609 ProAspPheLeuSerGlyIleValGlyLysIleAspSerLysSerIlePheCys 1628
Db |||||
349 CCAGATCCTGTGAATGGCATGGTGTGATCAAA----- 384
QY 1629 SerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeu 1648
Db |||||
385 -----GGCATC 390
QY 1649 LysProGlySerLysValLeuLeuPheCysAspProGlyPheGlnLeuValGlyAsnPro 1668
Db |||||
391 CACTTCGGATCCCAATTAATATCTTGTACTAAAGATACCGACTCATTTGTTCTCTCG 450
QY 1669 valGlnTyrCysLeuAsnGlnGlyGln-----TrpThrGlnProLeuProHisCys 1685
Db |||||
451 TCTGCCACATGCATCATCTCAGGTGATCTGTCATTTGGGATAATGAACACCTATTGT 510
QY 1686 GluArgIleSerCysGlyValProProProLeuGluAsnGly-----PheHisSerAla--- 1703
Db |||||
511 GACAGAAATTCCTTGTGGGCTACCCCCACCATCACCATTGGAGATTTCATTAGCACCCAAC 570
QY 1704 ---AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGly----- 1719
Db |||||
571 AGAGAGAAATTTTCACTATGATCGATCAGTGTGACCTACCGTGCATCTCTGGAGCGGAGG 630
QY 1720 -----TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsn----- 1733
Db |||||
631 AGAAAGTGTGAGCTTGTGGGTGAGCCCTCCATATATCTGCACCCAGCAATGACGATCAA 690
QY 1734 ---GlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaValGly 1752
Db |||||
691 GTGGGCATCTGGAGCGGCCCCCTCTAGTGCATTATATACCTAACAAATGC----- 741
QY 1753 SerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCys 1772
Db |||||
741 ----- 741
QY 1773 ValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysCysLysAlaPro 1792
Db |||||
742 -----ACGCCT 747
QY 1793 GlyAsnProGluAsnGly-----HisSerSerGlyGluIleTyrThrValGlyAla 1809
Db |||||
748 CCAATGTGGAATATGATATTTGGTATCTGACACAGAGAGCTTATTTTCTTTAAATGAA 807
QY 1810 AlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIleThrCys 1829
Db |||||
808 GTTGTGGAGTTTAGGTGTGAGCTTGTGTCATGAAGAGACCCCGCGGTGAGTGC 867
QY 1830 LeuGluSerGlyGluTrpAsnHisLeuIleProTyrCysLysAlaValSerCysGlyLys 1849
Db |||||

Db 868 CAGGCCCTGAAACAAATGGGAGCCGAGCTACCAAGCTGCTCCAGGGTANGTCAGCCACT 927
Qy 1850 ProAlaIle-----ProGluAsnGlyCysIleGluGluLeuAlaPheThrPheGlySer 1867
Db 928 CCAGATGCTCTGCATGCTGAGGTATCCCAAGGGACAGGCAACTTTTCACTTGGGCAG 987
Qy 1868 LysValThrTyArgCysAsnLysGlyTyThrLeuAlaGlyAspLysGluSerCys 1887
Db 988 GAAAGTGTCTACAGCTGTGAGCCGCTACGACCTCAGAGGGGCTCGCTATGCGCTGC 1047
Qy 1888 LeuAlaAsnSerSerTyPheSerHisSerProProValCysGluProValLysCysSerSer 1907
Db 1048 ACACCCAGGGAGACTGGAGCCCTGAGCCCTGAGCCCTGAGTGAAGTGAATCTCTGATGAC 1107
Qy 1908 ---ProGluAsnIleAsnAsnGlyLysTyIleLeuSer---GlyLeuThrTyIleLeuSer 1925
Db 1108 TTCATGGGCCCAACTTCTTAATGGCGGTGCTGCTATTTCCAGTAAATCTCCAGCTTGGAGCA 1167
Qy 1926 ThrAlaSerTySerCysAspThrGlyTySerLeuGlnGlyProSerIleIleGluCys 1945
Db 1168 AAGTGGATTTGTGTGATGAAGCAATTCATTAATAAAGGAGCTCTGCTAGTTACTGT 1227
Qy 1946 ThrAlaSerGly-----IleTrpAspArgAlaProProAlaCysHisLeuValPhe 1962
Db 1228 GTCTTGGCTGGAAATGGAAAGCCCTTGAATAGCAGTGTTCAGTGTGTGAACAATCTTT 1287
Qy 1963 CysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsn-----Phe 1979
Db 1288 TGTCCAGTCTCTCAGTATTCCTAATGGGAGACACACAGGAAACCTCTGGAAGTCTTT 1347
Qy 1980 ThrPheArgAsnThrValThrTyThrCysLys-----GluGlyTy 1993
Db 1348 CCCTTTGGGAAACAGTAAATACATGCGACCTGCGACCCCAACCCAGACAGAGGAGCTTC 1407
Qy 1994 ThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAsp 2013
Db 1408 GACCTCATGGAGAGACCACTCCCTGCTCAGAGTGAC----- 1446
Qy 2014 GlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerProGlu 2033
Db 1447 -----CCTCAA 1452
Qy 2034 ThrAlaHisArgLeuPheGlyAspIleAlaPheTyThrCysSerAspGlyTySerLeu 2053
Db 1452 ----- 1452
Qy 2054 AlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGlyGln 2073
Db 1453 -----GGAAATGGGGTTTGGAGCAGCCCT----- 1476
Qy 2074 AspMetProArgCys---IleAlaHisPheCysGluLysPheProProSerValSerTySer 2092
Db 1477 ---GCCCTCGCTGTGAATTCGGGTCACTGTCAAGCCCAAGATCAATTTCTGTGTC 1533
Qy 2093 IleLeuGluSerValSerLysAlaLys---PheAlaAlaGlySerValValSerPheLys 2111
Db 1534 AAGTTGAAACCAACCAATGCATCTGACTTTCCCATTTGGGACATCTTTAAAGTACGAA 1593
Qy 2112 CysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGln 2131
Db 1594 TCCGCTCTGATGAC---TACGGAGCCATCTCTATACATGCTAGATTAACCTGGTC 1650
Qy 2132 TrpAsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSer 2151
Db 1651 TGGTCA---AGTCCCAAGATGTC---TGTAAACGTAAATCATGTAAACTCTCTCAGAT 1704
Qy 2152 IleMetAsnGlyTyThrAlaSerGly---SerAsnTySerPheGlyAlaMetValAlaTy 2170
Db 1705 CCAGTGAATGGCATGGTGTGATGATCAGACATCCAGGTTGGATCCAGAAATCAACTAT 1764
Qy 2171 SerCysAsnLysGlyPheTyIleLysGlyGluLysLysSerThrCysGluAlaThrGly 2190
Db 1765 TCTTGTATACAGGGGACCGACTCATTTGTTGCTCATCTGCTGAATGATCTCTCTGGGC 1824

Qy 2191 -----GlnTrpSerSerProIleProThrCysHisProValSerCysGlyGluPro 2207
Db 1825 AATGCTGCCCATTTGGAGCAGCAAGCCCAATTTGTCAACGAATTTCTTGTGGGCTACCC 1884
Qy 2208 ProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGlu-----Ser 2225
Db 1895 CCCACCATCGCCAAATGGAGATTTCAATTAGCAACAACAGAGAGAAATTTTCACTATGATCA 1944
Qy 2226 GluValArgTyGlnCysAsnProGly-----TyrLysSerValGly 2239
Db 1945 GTGGTGACCTACCGCTGCAATCCCGGAAGCGGAGGAGGAAGGTGTTTTCAGCTTGTGGGT 2004
Qy 2240 SerProValPheValCysGlnAlaAsnArgHis-----TrpHisSerGluSer 2255
Db 2005 GAGCCCTCCATATACTGCACAGCAATGACGATCAAGTGGGCATCTGAGCGCCCGGCC 2064
Qy 2256 ProLeuMetCysValProLeuAspCysGlyLysProProProIleGlnAsnGlyPheMet 2275
Db 2065 CCTCAGTCATTATACCTAACAAATGC---AGCCTCCAAATGTGGAAATGGAAATATG 2121
Qy 2276 LysGlyGluAsn-----PheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
Db 2122 GTATCTGACAACAGAGCTTATTTCTTAAATGAATGTTGTGGAGTTTAGGTGTGAGCCT 2181
Qy 2293 GlyTyGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLys 2312
Db 2182 GGTCTTGTGATGAAGGAGCCCGCGCTGTGAAGTGCAGGCCCTGACCAAAATGGAGCCG 2241
Qy 2313 LysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGln 2332
Db 2242 GAGCTA---CCAGCTGCTCCAGGGTATGTCAGCCACCTCCAGATGCTCTGATGCTGAG 2298
Qy 2333 LeuValLeu-----LysGluLeuThrThrGluValGlyValValThrPheSerCysLys 2350
Db 2299 CGTACCCCAAGGACAGGACAACTTTTCCACCGGCGCAGGAGTGTCTACAGCTGTGAG 2358
Qy 2351 GluGlyHisValLeuGlnGlyProSerValLysCysLeuProSerGlnGlnTrpAsn 2370
Db 2359 CCGGCTATGACTCAGAGGGCTGCGTCTATGCGCTGACACCCCGGAGAGACTGAGC 2418
Qy 2371 AspSerPheProValCysLysIleValLeuCysThrPro-----ProProLeuLysSer 2388
Db 2419 CTGACGCCCCACATGTGAAGTGAATCTGTGATGACTTCATGGGCCCAACTTCTTAAT 2478
Qy 2389 PheGlyValProIleProSerSerAlaLeuHisPheGlySerThrValLysTySerCys 2408
Db 2479 GCGCGTGTCTATTTCAGTAAAT---CTCCAGCTTGGAGCAAAAGTGGATTTTGTGT 2535
Qy 2409 ValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGly----- 2426
Db 2536 GATGAAGGATTTCAATTAAGGCGAGCTCTGCTAGTATTGTGTCTTGGCTGGAATGGAA 2595
Qy 2427 ---ThrTrpSerSerProLeuProGluCysValProValGluCysProGlnProGluGlu 2445
Db 2596 AGCCTTTGAATAGCAGTGTTCAGTGTGTGAACAATCTTTTGTCCAGTCTCCAGATT 2655
Qy 2446 IleProAsnGlyIle-----IleAspValGlnGlyLeuAlaTyLeuSerThrAla 2462
Db 2656 ATTCCTAATGGGAGACACACAGGAAACCTCTGGAAGTCTTTCCCTTTGGAAAAGCAGTA 2715
Qy 2463 LeuTyThrCysLysPro-----GlyPheGluLeuValGlyAsnThr 2476
Db 2716 AATTACATGAGGACCCCCCAGACAGAGGAGCAGCTTTCACCTCATTTGAGAGAGC 2775
Qy 2477 ThrThrLeuCysGlyGlu-----AsnGlyHisTrpLeuGlyGlyLysProThr 2492
Db 2776 ACCATCGCTGCACAGTGCCTCAAGGAATGGGTTTGGAGCAGCCCTGCGCCTCGC 2835
Qy 2493 CysLysAlaIle---GluCysLeuLysProLysGluIleLeuAsnGlyLysPhe----- 2509
Db 2836 TGTGAATTTCTGGTCACTGTCAAGCCCGCAGATCATTTTCTGTTGCAAGTGTGAAAC 2895

Db 4900 ATG---GTAGGTCCACACTGTGTCAGTCCAGACCAATGTCAGATGGGGCCCAAGCTG 4956
Qy 3174 lIeSerCysSerProLysLysCysProLeuProGluAsnIle-----ThrHisIleLeu 3191
Db 4957 CCACACTGCTCCAGGTGTGTCCAGCGCTCCAGAAATCTCTCATGTGTGATACCTTA 5016
Qy 3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
Db 5017 AGCCATCATGAGCAACTTTTCACCTGGCAGGAAGTGTCTTACAGCTGTGAGCCAGCTAT 5076
Qy 3212 ThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTyrGluProPhe 3231
Db 5077 GACCTCAGAGGGCTGCGTCTCTGCACTGCAGCCCGCAGGAGACTGGAGCCCT----- 5130
Qy 3232 SerAspGluSerCysSerProValSerCysGlyLys-----ProGluSerProGluHis 3249
Db 5131 GAAGCCCTTAGATGATACAGTAATCTCTGTGATGACTTCTCTGGCCCAACTCCCT---CAT 5187
Qy 3250 GlyPheVal-----ValGlySerLysTyrThrPheGluSerThr 3262
Db 5188 GCGCGTGTCTACTTCCACTTAATCTCCAGCTTGGGCGCAAGGTGTCTCTT----- 5238
Qy 3263 lIeIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys--- 3281
Db 5239 -----GTTTGGATGAGGGTTCCGATTAAGGCGAGGTCTCTAGTCAATTGTGTCT 5289
Qy 3282 -----GlnGluAsnArgGlnTyrPheSerGlyGlyValAlaIleCysLysGluThrArgCys 3299
Db 5290 TTGGCTGGAATGAAGCCCTTTTGGATAGCAGTGTTCAGTGTGTGAACAATCTTTTGT 5349
Qy 3300 GluThrProLeuGluPheLeuAsnGlyLys-----AlaAspIleGlu 3313
Db 5350 CCAAAATCTCTCAGCTATCTTAAATGGAGACACACAGGAACCTCTTTTGGAGATATT--- 5406
Qy 3314 AsnArgThrThrGlyProAsnValValTyrSerCys-----AsnArgGly--- 3328
Db 5407 -----CCTTGTGAAAGAAATATCTTACGATCGGACACCCACCCAGCAGAGGATG 5460
Qy 3329 ---TyrSerLeuGluGlyProSerGluAlaHisCysThrGlu-----AsnGly 3343
Db 5461 ACCTTCAACTCATTTGGGAGAGCTCCATCGCTGACACAGTGACCTCAAGGGAATGG 5520
Qy 3344 ThrTyrSerHisProValProLeuCysLys-----ProAsnProCysProValPro 3360
Db 5521 GTTTGGAGCAGCCCTGCGCTCGCTGTGAATTTCTGTCTCTGCTGCTGCTGCTGCTGCT 5580
Qy 3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
Db 5581 CCAAGATC-----CAAAACGGCGCATTTACATTTGAGGACACGTA 5619
Qy 3381 SerIle-----LysCysArgGluGlyPheLeuLeuGln 3391
Db 5620 TCTCTATATCTTCTGGGATGACAAATCAGTACACTGTGTGACCCCGCTACCTGTATGTG 5679
Qy 3392 GlyHisGlyIleIleThrCysAsnProAspGluThrTyrThrGlnThrSerAlaLysCys 3411
Db 5680 GGAAGAGGGCTTCATTTCTGTACAGACCAGGGAATCTGAGGCAATTTGATCATTTATTC 5739
Qy 3412 GluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHis 3431
Db 5740 AAGAAGTAATTTAGCTTCCACCTGTTATG---AATGGATCTGAGGAGGTAGAA 5796
Qy 3432 -----TyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMet 3447
Db 5797 ATGAAAAAGTATATCACTATGAGATTATGTGACTTTGACGTGTGACATGGGTATATCT 5856
Qy 3448 LeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTyrThrSerProPheIle 3467
Db 5857 CTGGAAGGAGCTCCCTGGAGCAGTCCAGCGGATGACAGATGG---GACCCCTCTCTCTG 5913
Qy 3468 CysArgAlaValCysArg 3473
Db 5914 GCCAAATGTACTCTCTCT 5931

RESULT 5

US-09-023-655-1265
; Sequence 1265, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1265:

SEQUENCE CHARACTERISTICS:

LENGTH: 6951 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g30185

US-09-023-655-1265

Alignment Scores:

Pred. No.:	3,77e-157	Length:	6951
Score:	2027.00	Matches:	563
Percent Similarity:	40.08%	Conservative:	265
Best Local Similarity:	27.25%	Mismatches:	832
Query Match:	10.15%	Indels:	406
DB:	4	Gaps:	85

US-09-977-053-4 (1-3571) x US-09-023-655-1265 (1-6951)

Qy	1609	ProAspPheLeuSerGlyIleValGlyLysValIleAspSerLysSerIlePheCys	1628
Db	349	CCAGATCTCTGAATGGCATGTGTCATCAAA-----	384
Qy	1629	SerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeu	1648
Db	385	-----GGCATC 390	
Qy	1649	LysProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnPro	1668
Db	391	CAGTTGGATGCCAAATTAATTTCTTGTACTAAAGGATACCGACTCATTTGTTCTCTCG	450
Qy	1669	ValGlnTyrCysLeuAsnGlnGlyGln-----TipThrGlnProLeuProHisCys	1685

451	DB	TCGCCACATGCAATCATCTCAGGTGATACTGTCATTTGGGATTAATGAAACACCTATTGTG	510
1686	QY	GLuArgIleSerCysGlyValProProLeuGluAsnGly---PheHisSerAla---	1703
511	DB	GACAGAAATCTCTGTGGGCTACCCCCACCATCACCAATGAGGANTTTCATTAGCACCAAC	570
1704	QY	---AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGly---	1719
571	DB	AGAGAGAAATTTTCACTATGATCAGTGTGACCTACCGCTGCAATCTCGAAGCGGAGG	630
1720	QY	-----TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsn---	1733
631	DB	AGAAAGGTGTTTCAGCTGTGTGGGTGAGCCCTCCATATACTGCCACCAATGACGATCAA	690
1734	QY	--GlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaValGly	1752
691	DB	GTGGGCATCTGGAGCGCCCCCGCTTCAGTGCATATACCTAACAAATGC-----	741
1753	QY	SerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyrIleCysSerCys	1772
741	DB	-----	741
1773	QY	ValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysCysLysAlaPro	1792
742	DB	-----AGGCCT	747
1793	QY	GlyAsnProGluAsnGly-----HisSerSerGlyGluIleTyrThrValGlyAla	1809
748	DB	CCAAATGTGAAATGGAATATTGGTATCTGACAAACAGAAAGCTTATTTCTTAAATGAA	807
1810	QY	AlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIleThrCys	1829
808	DB	GTGTGTGAGTTTAGTGTGAGCTGGCTTTCTCATGAAGAGACCCCGCTGTGAAGTGC	867
1830	QY	LeuGluSerGlyGluTrpAsnHisLeuIleProTyrCysIysAlaValSerCysGlyLys	1849
868	DB	CAGGCCCTGAACAAATGGAGCCGGAGCTACCAAGCTGCTCCAGGGTATGTCAGCCACCT	927
1850	QY	ProAlaIle-----ProGluAsnGlyCysIleGluGluLeuAlaPheThrPheGlySer	1867
928	DB	CCAGATCTCTGCATGTCGAGCGTACCACCAAGGGACCAAGGACAACTTTTTCACCTGGGCAG	987
1868	QY	LysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLysGluSerSerCys	1887
988	DB	GAAGTGTCTACAGCTGTGAGCCCGGCTACGACCTCAGAGGGGCTGGTCTATCCGGTGC	1047
1888	QY	LeuAlaAsnSerSerTrpSerHisSerProProValCysGluProValLysCysSerSer	1907
1048	DB	ACACCCAGGGAGACTGAGCCCTGCAGCCCCACATGTGAAGTGAATCTCTGTGATGAC	1107
1908	QY	-----ProGluAsnIleAsnAsnGlyLysTyrIleLeuSer---GlyLeuThrTyrLeuSer	1925
1108	DB	TTCATGGCCCAACTCTTAATGGCCGTGTGCTATTTCCAGTAAATCTCCAGCTTGGAGCA	1167
1926	QY	ThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIleGluCys	1945
1168	DB	AAAGTGAATTTGTTGTGATGAGGATTTCAATTAAGGCGAGCTGTGCTAGTTACTGT	1227
1946	QY	ThrAlaSerGly-----IleTrpAspArgAlaProProLacCysHisLeuValPhe	1962
1228	DB	GTCTTGCTGGAAAGGAAAGCCTTTGGAATAGCAGTGTTCAGTGTGTGAACAAATCTTT	1287
1963	QY	CysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsn-----Phe	1979
1288	DB	TGTCCAAGTCTCCAGTTATTCCTTAATGGGAGACACACAGAAACCTCTGGAGTCTTT	1347
1980	QY	ThrPheArgAsnThrValThrTyrThrCysLys-----GluGlyTyr	1993
1348	DB	CCCTTTGGAAAGCAGTAATTAACATGCGACCCGCCACCCAGACGAGGGACGAGCTTC	1407
1994	QY	ThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAsp	2013

1408	Db	GACCTCATTTGGAGAGACCACTCCGCTGCACAGTGC-	
2014	Qy	GlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerProGlu	2033
1447	Db	-----	1452
2034	Qy	ThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeu	2053
1452	Db	-----	1452
2054	Qy	AlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGlyGln	2073
1453	Db	-----GGGAATGGGGTTTGGAGCAGCCCT	1476
2074	Qy	AspMetProArgCys---IleAlaHisPheCysGluLysProProSerValSerTyrSer	2092
1477	Db	---GCOCCTCGCTGTGGAAATCTTGGGTCACTGTCAAGCCCCCAGATCATTTCTTGTTGCC	1533
2093	Qy	IleLeuGluSerValSerLysAlaLys---PheAlaAlaGlySerValValSerPheLys	2111
1534	Db	AAGTTGAHAACCCAAACCAATGCATCTGACTTTCCTATGGGACATCTTTAAAGTAGCAA	1593
2112	Qy	CysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGln	2131
1594	Db	TGCGTCTCGATGAC---TACGGGAGGCCATTCTCTATCATGTCTAGATAACCTGGTC	1650
2132	Qy	TrpAsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSer	2151
1651	Db	TGTTGCA---AGTCCCCAAAGATGC---TGTAACGGTAAATCATGTAAACTCTCTCCAGAT	1704
2152	Qy	IleMetAsnGlyTyrAlaSerGly---SerAsnTyrSerPheGlyAlaMetValAlaTyr	2170
1705	Db	CCAGTGAATGGCATGGTGTCATGTATCACAGACATCCAGGTTGGATCCAGAATCAACATAT	1764
2171	Qy	SerCysAsnLysGlyPheTyrIleLysGlyGluLysLysSerThrCysGluAlaThrGly	2190
1765	Db	TCTTGTACTACAGGGCACCGCACTCATTTGGTCACATCTGCTGAAATGATCTCTCTCGGGC	1824
2191	Qy	-----GlnTrpSerProIleProThrCysHisProValSerCysGlyGluPro	2207
1825	Db	ANTGCTGCCATTGGAGCAGACGCCCAATTTGTTCACGAATCTCTTGTGGGTACCC	1884
2208	Qy	ProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGlu-----Ser	2225
1885	Db	CCCACTATGCCCAATGAGATTTTCATTAGCACCAACAGAGAGAATTTTCACTATGGATCA	1944
2226	Qy	GluValArgTyrGlnCysAsnProGly-----TyrLysSerValGly	2239
1945	Db	GTGGTGACCTTACCCTGCTCAATCTCTGAGCGGAGGAGAAAGGTGTTTGTAGCTTGCGGT	2004
2240	Qy	SerProValPheValCysGlnAlaAsnArgHis-----TrpHisSerGluSer	2255
2005	Db	GAGCCCTCCATATACTGCACCACGAATGACATCAAGTGGGCATCTGGAGCGCCCGGCC	2064
2256	Qy	ProLeuMetCysValProLeuAspCysGlyLysProProProIleGlnAsnGlyPheMet	2275
2065	Db	CCTCAGTGCATTATACCTAACAAATGC---ACGCCCTCCAAATGTGGAAAATGAATAATTG	2121
2276	Qy	LysGlyGluAsn-----PheGluValGlySerLysValGlnPhePheCysAsnGlu	2292
2122	Db	GTATCTGACAAACAGAGCTTATTTTCTTAAATGAAGTTGTGGAGTTTAGGTGTGACCT	2181
2293	Qy	GlyTyrGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLys	2312
2182	Db	GGCTTTGTCTATGAAGAGACCCCGCTGTGAACTGCCAGGCCCTTGAAACAATGGGAGCCG	2241
2313	Qy	LysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsnGln	2332
2242	Db	GAGCTA---CCAAGTGTCTCAGGGTATGTGACCCACTTCAGATGTCTGCTGCTGAG	2298
2333	Qy	LeuValLeu-----LysGluLeuThrThrGluValGlyValValThrPheSerCysLys	2350
2299	Db	CGTACCACCAAGGACCAAGCAACTTTTCAACCGGGCAGGAAGTTTCTACAGCTGTGAG	2358

Db 4360 TGTCTAAGTCTTGAAGCAACTGTAGACGAAATCATGTGGACCTCCACAGAACCCCTTC 4419
QY 3015 ---GlyThrValAsn---GlyThrAspPheAspCysGlyIysAlaAlaArgIleGlnCys 3032
Db 4420 AATGGATGGTGCATATAAACAACAGATACACAGTTTGGATCAACAGTTAATTATTCTGT 4479
QY 3033 PheIysGlyPheIysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln--- 3051
Db 4480 AATGAAGGGTTTCGACTCATTTGGTTCCTCCCATCTACTACTTGTCTCGCTCAGGCAATPAT 4539
QY 3052 ---TyrSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMet 3069
Db 4540 GTACATGGGATGAAGAGCCACTTATTGTGAGATCATATCTTGTGAGCCACCTCCAAC 4599
QY 3070 IleProAsnAla---PheIleSer---GluThrSerSerTyrIysGluAsnValIle 3086
Db 4600 ATATCCATGGAGACTTCTACAGCAACAATAGAACATCTTTTCAATGGAACGGTGGTA 4659
QY 3087 ThrTyrSerCysArgSerGly---TyrValIleGlnGlySerSer 3100
Db 4660 ACTTACCAGTGCACACTGACACACAGATGAGAACAGCTGTTTGAGCTTGTGGGAGAACGG 4719
QY 3101 AspLeuIleCysThrGluIys---GlyValTyrSerGlnProTyrProVal 3116
Db 4720 TCAATATATTGCACACAGCAAGATGATCAAGTTGGTGTGTTGGAGCAGCTCCCTCCG 4779
QY 3117 CysGluProLeuSerCysGlySerProProSerValAlaAsnAlaAlaThrGlyGlu 3136
Db 4780 TGTATTCTACTAATAAATGACAGCTCCAGAGTTCGAAATGCAATTAGATACACAGGA 4839
QY 3137 AlaHisThrTyrGluSer---GluValIysLeuArgCysLeuGluGlyTyrThr 3153
Db 4840 AACAGGAGTTCTTTTCCCTCACTGAGATCATGATTTAGATGTCAGCCGGGTTTGT 4899
QY 3154 MetAspThrAspThrAspThrPheThrCysGlnIysAspGlyArgTyrPheProGluArg 3173
Db 4900 ATG---GTAGGTCCACACACTGTCAGTGCAGACCAATGCGAGATGGGGCCCAAGCTG 4956
QY 3174 IleSerCysSerProIysCysPheProLeuProGluAsnIle---ThrHisIleLeu 3191
Db 4957 CCACACTGCTCCAGGGTGTGTCAGCCGCTCCAGAAATCTCGCATGTGAGCAGTACCCTA 5016
QY 3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
Db 5017 AGCCATCAGGACAACTTTTCACTGGCAGGAAGTGTCTACAGCTGTGAGCCAGCTAT 5076
QY 3212 ThrPheGluGlyValaenIleSerValCysGlnLeuAspGlyThrTyrProProPhe 3231
Db 5077 GACTCAGAGGGGTGCTGCTCTGCACTGCACGCCCCAGGAGACTGAGCCCT--- 5130
QY 3232 SerAspGluSerCysSerProValSerCysGlyIys---ProGluSerProGluHis 3249
Db 5131 GAAGCCCTAGATGTACAGTGAATCTGTGATGACTTCTCGGCCCAACTCCCT---CAT 5187
QY 3250 GlyPheVal---valGlySerIysTyrThrPheGluSerThr 3262
Db 5188 GGCGGTGTGCTACTTCCACTTAATCTCCAGCTGGGGCAAGGTGTCTTT--- 5238
QY 3263 IleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys--- 3281
Db 5239 ---GTTTCGATGAAGGTTCCGATTAAGGCGAGGTCTGCTAGTCATTGTGTC 5289
QY 3282 ---GlnGluAsnArgGlnTyrSerGlyGlyValAlaIleCysIysGluThrArgCys 3299
Db 5290 TTGGCTGGAATGAAGCCCTTTTGAATGAGCAGTGTCCAGTGTGTGAACAAATCTTTTGT 5349
QY 3300 GluThrProLeuGluPheLeuAsnGlyIys---AlaAspIleGlu 3313
Db 5350 CCAAACTCTCCAGCTATCTTAATGGGAGAGACACAGGAACCTCCCTTTGGAGATATT--- 5406
QY 3314 AsnArgThrThrGlyProAsnValValTyrSerCys---AsnArgGly--- 3328

Db 5407 ---CCCTATGCAAAAGAAATATCTTACGCATGCGACACCCACCCACAGACAGAGGATG 5460
QY 3329 ---TyrSerLeuGluGlyProSerGluAlaHisCysThrGlu---AsnGly 3343
Db 5461 ACCTTCAACCTCATTTGGGAGAGCTCCATCGCTGCACAACTGACCCCTCAAGGGGAATGG 5520
QY 3344 ThrTyrSerHisProValProLeuCysIys---ProAsnProCysProValPro 3360
Db 5521 GTTTGGAGAGCCCTGCGCTCGCTGGAACCTTCTGCTCTGCTGCTGCCCATCA 5580
QY 3361 PheValIleProGluAsnAlaLeuLeuSerGluIysGluPheTyrValAspGluAsnVal 3380
Db 5581 CCCAAGATC---CAAAACGGGCAATTACATTGGAGGACACGTA 5619
QY 3381 SerIle---LysCysArgGluGlyPheLeuLeuGln 3391
Db 5620 TCTCTATATCTTCTGGGATGACAATCAGCTTGTGACCCCGCTACTCTGTTAGT 5679
QY 3392 GlyHisGlyIleIleThrCysAsnProAspGluThrTyrThrSerAlaIysCys 3411
Db 5680 GGAAGGGCTCATTTCTGTACAGACAGGAGATCTGGACCAATTGGATCATTTATTC 5739
QY 3412 GluIysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHis 3431
Db 5740 AAGAAGTAATGTAGCTTCCCACTGTTATG---AATGGAATCTCGAAGGAGTTAGAA 5796
QY 3432 ---TyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMet 3447
Db 5797 ATGAAAAAGATATATCATCTATGAGATTATGTGACTTTTGAAGTGAAGATGGGTATAT 5856
QY 3448 LeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTyrThrSerProProIle 3467
Db 5857 CTGGAGGAGTCCCTGGAGCAGTGCAGGCGAGTGCAGGCGATGCACATG---GACCTCTCTG 5913
QY 3468 CysArgAlaValCysArg 3473
Db 5914 GCCAAATGATACCTCTCGT 5931
RESULT 6
5256642-1
; Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
; RECEPTOR 1 (CR1) AND A THROMBOTIC AGENT, AND THE METHODS OF
; USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO: 1
; LENGTH: 6951
5256642-1
Alignment Scores:
Pred. No.: 3,09e-156 Length: 6951
Score: 2016.00 Matches: 562
Percent Similarity: 39.98% Conservative: 264
Best Local Similarity: 27.20% Mismatches: 834
Query Match: 10.09% Indels: 406
DB: 6 Gaps: 85
US-09-977-053-4 (1-3571) x 5256642-1 (1-6951)
QY 1609 ProAspPheLeuSerGlyIleValGlyIysValIysIleAspSerLysSerIlePheCys 1628

Db 4189 GAACCTTCTGTCGTGGTCACTGTAAACCCAGAGCAGTTTCCATTGTCAGTCCCT 4248
Qy 2961 -----AsnGlyPheSerPheIleHisGlyGlyHieIleGlnThrClnCysPhe 2976
Db 4249 ACGATCCCAATTAATGACTTTGAGTTCCAGTCGGGACATCTTGAATATGAATGCGGT 4308
Qy 2977 ProGlyTyr-----LysLeuHisGlyAsnSerSerArgCysLeuSerAsnGlySer 2994
Db 4309 CTGGGTATTTTGGGAAATGTTCTCTATCTCC-----TGCCTAGAAATCTGGTC 4359
Qy 2995 TrpSerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGlnTyr 3014
Db 4360 TGGTCAAGTGTGAAGACAACATGTAGACGAAATCATGTGGACCTCCACCGAACCCCTTC 4419
Qy 3015 ---GlyThrValAsn---GlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCys 3032
Db 4420 AATGGAAATGGTGCATATAAACAACATACACAGTTTGGATCAACAGTTAATTTCTTGT 4479
Qy 3033 PheLysGlyPheLysLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln--- 3051
Db 4480 AATGAAGGTTTTCGACTCAATGGTTCCCATCTACTTGTCTGCTCTCGAGCAATAAT 4539
Qy 3052 -----TrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMet 3069
Db 4540 GTCACATGGGATAAGAGGCACTTATTTGTGAGATCATATCTTGTGAGCCACTCCAACC 4599
Qy 3070 IleProAsnAla---PheIleSer-----GluThrSerSerTrpLysGluAsnValIle 3086
Db 4600 ATATCAATGGAGACTTCTACAGCAACAATAGAACATCTTTTACAAATGGAAGGTGTA 4659
Qy 3087 ThrTyrSerCysArgSerGly-----TyrValIleGlnGlySerSer 3100
Db 4660 ACTTACCAGTTCACACTGGACAGATGAGAACAGCTGTTTGAGCTTGTGGGAGAACGG 4719
Qy 3101 AspLeuIleCysThrGluLys-----GlyValTrpSerGlnProTrpVal 3116
Db 4720 TCAATATATGTCACAGCAAGATGATCAAGTTGGTGTGTGGAGCGCCCTCCCTCGG 4779
Qy 3117 CysGluProLeuSerCysGlySerProProSerValAlaAsnAlaAlaThrGlyGlu 3136
Db 4780 TGTATTTCTACTAATAAATGACAGCTCCAGAGTTGAAATGCAATTAGATGACGGA 4839
Qy 3137 AlaHisThrTyrGluSer-----GluValLysLeuArgCysLeuGlyTyrThr 3153
Db 4840 AACAGGAGTTCTTTTCCCTCACTGAGATCATCAGATTTAGATGTGAGCCCGGTTTGT 4899
Qy 3154 MetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluArg 3173
Db 4900 ATG-----GTAGGTCCTCCACACTGTGACGTGCCAGACCAATGGCAGATGGGGCCGAGCTG 4956
Qy 3174 IleSerCysSerProLysLysCysProLeuProGluAsnIle-----ThrHisIleLeu 3191
Db 4957 CCACACTGCTCCAGGTGTGTGTCAGCGCTCCAGAAATCCTGCTGAGCATACCCCTA 5016
Qy 3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
Db 5017 AGCCATCAGGACAACACTTTTCCCTGGCAGGAAGTGTCTACAGCTGTGAGCCGACGAT 5076
Qy 3212 ThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPhe 3231
Db 5077 GACCTCAGAGGGGTGCGTCTCTGCACTGACGCCGCCAGGAGACTGGAGCCCT----- 5130
Qy 3232 SerAspGluSerCysSerProValSerCysGlyLys-----ProGluSerProGluHis 3249
Db 5131 GAAGCCCTTAGATGTACAGTGAATCTGTGATGACTTCTGGGCCAACTCCCT---CAT 5187
Qy 3250 GlyPheVal-----ValGlySerLysTyrThrPheGluSerThr 3262
Db 5188 GGCCGTGTGCTACTTCCACTTAATCTCCAGTCTGGGGCAAGGTGTCTCTT----- 5238
Qy 3263 IleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys--- 3281

Db 5239 -----GTTTGGATGAAGGGTTCGATTAAAGGCGAGTCTCTAGTCAATTGTC 5289
Qy 3282 -----GlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCys 3299
Db 5290 TTGGCTGGAATGAAGCCCTTTGGAATAGCAGTGTTCAGTGTGTGAACAATCTTTTGT 5349
Qy 3300 GluThrProLeuGluPheLeuAsnGlyLys-----AlaAspIleGlu 3313
Db 5350 CCAATCTCCAGCTATCTTAATGGAGACACACAGGAACTCCCTTTGGAGATAT--- 5406
Qy 3314 AsnArgThrThrGlyProAsnValValTyrSerCys-----AsnArgGly--- 3328
Db 5407 -----CCATATGGAAGAAATACTTACGATCGACACCCACCCACAGAGGGATG 5460
Qy 3329 ---TyrSerLeuGluGlyProSerGluAlaHisCysThrGlu-----AsnGly 3343
Db 5461 ACCTTCAACCTCATTTGGGGAGAGCTCCATCGCTGCACAAAGTGACCCCTCAAGGAAATGG 5520
Qy 3344 ThrTrpSerHisProValProLeuCysLys-----ProAsnProCysProValPro 3360
Db 5521 GTTTGGAGCAGCCCTGCCCCCTCGCTGTGAACCTTCTGTCTGCTGCTGCCCTGCCACATCCA 5580
Qy 3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
Db 5581 CCCAAGATC-----CAAAACGGGCAATTACATTGGAGGACACGTA 5619
Qy 3381 SerIle-----LysCysArgGluGlyPheLeuLeuGln 3391
Db 5620 TCTCTATATCTTCTGGGATGACATCAGCTACACTTGTGACCCCGCTACCTGTTAGTG 5679
Qy 3392 GlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCys 3411
Db 5680 GGAAGGGCTTCTTTCTGTACACACCGAGGAATCTGGAGCCAAATGGATCATTTATG 5739
Qy 3412 GluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHis 3431
Db 5740 AAAGAAGTAAATTTAGCTTCCACTGTTTATG---AATGGAATCTCGAAGGAGTAGAA 5796
Qy 3432 -----TyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMet 3447
Db 5797 ATGAAAAAAGTATATCATCTATGAGATTTATGTGACTTTGAAGTGTGAAGATGGGTATCT 5856
Qy 3448 LeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProPheIle 3467
Db 5857 CTGGAAGGAGTCTCCCTGGAGCCAGTCCAGGCGGATGACAGATGG---GACCTCTCTCTG 5913
Qy 3468 CysArgAlaValCysArg 3473
Db 5914 GCCAAATGTACTCTCTGT 5931
RESULT 7
5472939-1
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE M.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588,128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:1:
; LENGTH: 6951

Db 1945 GTGGTGACCTACCGCTGCAATCTCTGGAAAGCGAGGAGAAAGGTGTTTGTGAGCTTTGTGGGT 2004
Qy
Db 2240 SerProValPheValCysGlnAlaAsnArgHis-----TTPHisSerGluSer 2255
Db 2005 GAGCCCTCCATATCTGACACGACATGACGATCAGTGGGCATCTGAGCGGCCCGGCC 2064
Qy 2256 ProLeuMetCysValProLeuAspCysGlyLysProProLeuGlnAsnGlyPheMet 2275
Db 2065 CCTCAGTGCATTATACCTAACAAATGC--ACGCCCTCAAAATGTGAAATGGAATATTG 2121
Qy 2276 LysGlyGluAsn-----PheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
Db 2122 GTATCTGCAACAGAGCTTATTTCTTAAATGAAGTTGTGAGTTAGTGTGAGCCT 2181
Qy 2293 GlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyrAsnLys 2312
Db 2182 GGCCTTGTGATGAAGGACCGCGCTGTGAAGTGCACAGCCCTGAAACAATGGGAGCG 2241
Qy 2313 LysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGlnGln 2332
Db 2242 GAGCTA---CCNAGCTGCTCCAGGGTATGTGAGCCACCTCCAGATGTCTCGATGCTGAG 2298
Qy 2333 LeuValLeu-----LysGluLeuThrThrGluValGlyValValThrPheSerCysLys 2350
Db 2299 CGTACCCAAAGGACAGAGCAACTTTTCCCGGGCAGGAGTGTCTACAGCTGTGAG 2358
Qy 2351 GluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyrAsn 2370
Db 2359 CCCGGCTATGACCTCAGAGGGGTGCGTCTATGCGCTGCACACCCAGGAGACTGGAGC 2418
Qy 2371 AspSerPheProValCysLysValLeuValLeuCysThrPro-----ProProLeuLeuSer 2388
Db 2419 CTGCGAGCCGCCACATGTGAAGTGAATCTCTGATGACTTCATGGGCCAACTTCTTAAT 2478
Qy 2389 PheGlyValProLeuProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCys 2408
Db 2479 GGCCGTGTGCTATTTCCAGTAAAT---CTCAGCTTGGAGCAAAAGTGAGTTTGTGTGT 2535
Qy 2409 ValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGly----- 2426
Db 2536 GATGAAGATTTCATTAAGGAGCAGCTCTGCTAGTTATGTGTCTTGGCTGGAATGAA 2595
Qy 2427 ---ThrTyrSerSerProLeuProGluCysValProValGluCysProGlnProGluGlu 2445
Db 2596 AGCCTTTGGAAATAGCAGTGTCTCAGTGTGTGAACAAATCTTTTGTCCAAGTCTCTCAGT 2655
Qy 2446 IleProAsnGlyIle-----IleAspValGlnGlyLeuAlaTyrLeuSerThrAla 2462
Db 2656 ATTTCTAATGGAGACACACAGGAAACCTCTGGAAGTCTTTCCCTTTTGGAAAGCAGTA 2715
Qy 2463 LeuTyrThrCysLysPro-----GlyPheGluLeuValGlyAsnThr 2476
Db 2716 AATTACATGGAGCCCCACCCACGACAGGAGGAGCTTCGACCTCATTTGGAGAGC 2775
Qy 2477 ThrThrLeuCysGlyGlu-----AsnGlyHisTyrLeuGlyGlyLysProThr 2492
Db 2776 ACCATCCGCTGCACAAAGTGCCTCAAGGAAATGGGGTTTGGAGCAGCCCTGCCCTCGC 2835
Qy 2493 CysLysAlaIle---GluCysLeuLysProLysGluIleLeuAsnGlyLysPhe----- 2509
Db 2836 TGTGGAATCTGGGTCACTGTCAAGCCCGACGATCATTTCTTGTGGCCAGTTGAAACC 2895
Qy 2510 -----SerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGly 2527
Db 2896 CAACCAATGATCTGACTTTCCCATTTGGACATCTTTAAAGTACGAATGCCGCTCTGAG 2955
Qy 2528 PheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTyrAspValAsp 2547
Db 2956 TAC---TACGGGAGGCCATTCTCTATCATCATGTCTAGATAACCTGGTCTGTGCAAGTCCC 3012
Qy 2548 AlaProSerCysAsnAlaIleHisCysAspSerProGlnProLeuGlnAsnGlyPheVal 2567
Db 3013 AAAGATGTCTGTAAACGTAATCATGTAAAACTCTCTCAGATCCAGTGCATGGCATGGTG 3072

Qy 2568 GluGly---AlaAspTyrSerTyrGlyAlaIleLeuTyrSerCysPheProGlyPhe 2586
Db 3073 CATGTGATCACAGACATCCAGGTGGATCCAGATCAACTATTCTTGTACTACAGGGCAC 3132
Qy 2587 GlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly-----TyrSer 2602
Db 3133 CGACTCATTGTGCTACTCTGCTGAATGTATCTCTCAGGCAATATGTCGCCCATTTGGAGC 3192
Qy 2603 SerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPheGly 2622
Db 3193 ACGAAGCCGCCAATTTGTCAAGGAATCTCTGTGGGCTACCCCAACCATGCCCAATGGA 3252
Qy 2623 AspCysThrLysLeuLysAspGlnGlyTyrPheGluGlnGluAspAspMetMetGlu 2642
Db 3253 GATTTCAAT----- 3261
Qy 2643 ValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTyrGlu 2662
Db 3261 ----- 3261
Qy 2663 AsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSer 2682
Db 3262 -----AGCACCAACAGAGAGATTTTCACTATGATCAGTGGTGACC 3303
Qy 2683 TyrThrCysAsnProGly-----TyrGluLeuLeuGlyAsnProVal 2696
Db 3304 TACCGCTGCAATCTTGGAAAGCAGAGGAGAAAGGTGTTTGTGAGCTTGTGGGTGAGCCCTCC 3363
Qy 2697 LeuIleCysGlnGluAsp-----GlyThrTyrAsnGlySerAlaProSerCys 2712
Db 3364 ATATACCTGCACAGCAATGACGATCAGTGGGCATCTGAGCGGCCGCCCTCTAGTGC 3423
Qy 2713 IleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGlu 2732
Db 3424 ATTATACCTAACAAATGCACGCTCCAAATCTGGAATGGAATATTGGTATCTGCAAC 3483
Qy 2733 ThrSerMetGlySer-----AlaValGlnTyrSerCysLysProGlyHisIleLeu 2749
Db 3484 AGAAGCTTATTTTCTTAAATGAAGTTGTGAGCTTTAGGTGTGCGCTGGCTTGTGCATG 3543
Qy 2750 AlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTyrTyrSerGlyAlaSerProArg 2769
Db 3544 AAAGAGCCCGCGCTGTGAGTGCAGGCCCTCGAACAAATGGAGGAGGAGTATCCCAAGC 3603
Qy 2770 CysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySer 2789
Db 3604 TGCTCCAGGGTG---TGTACGCGCGCTCCAGAAATCCTGCATGGTGAGCATACCCCAAGC 3660
Qy 2790 -----AsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrVal 2806
Db 3661 CATCAGGACAACTTTTACCTGGGCGAGGAAGTGTCTACAGCTGTGAGCTTGCTATGAC 3720
Qy 2807 LeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTyrAspGluAspGluPro 2826
Db 3721 CTCAGAGGGCGTGGCTCTCTGCATGACACCCAGGAGACTGGAGCCCTGAGGCCCG 3780
Qy 2827 IleCysIleProValAspCysSerSerProValSerAlaAsnGlyGlnValArgGly 2846
Db 3781 AGATGTGAGTGAATCCTGTGAT-----GACTTCTGGGTCAACTCCCTCAT 3828
Qy 2847 AspGluTyrThrPheGln-----LysGluIleGluTyrThrCysAsn 2860
Db 3829 GGCGGTGTGCTATTTTCCACTTAATCTCAGCTGGGCGCAAGGTGTCTTTGTCTGTAT 3888
Qy 2861 GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySer----- 2878
Db 3889 GAGGGTTTCCTTAAGGCGCAGTTCCTGTAGTCTTGTCTGTCTGGAATGAGAAGC 3948
Qy 2879 ---TyrSerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeu 2897
Db 3949 CTTTGAATAACAGTGTCTCTGTGTGNAACATATCTTTTGTCCAAATCTCTCCAGCTATC 4008

Qy	2998	Ala	Asn	Gly	Val	Thr	Glu	Gly	Leu	-----	Asp	Tyr	Gly	Phe	Met	Lys	Glu	Val	Thr	2914
Db	4009	CTT	AA	TGG	GAG	ACAC	ACAG	GAA	CTC	CTC	CTG	GAG	ATA	TTCC	TCT	TGG	AGAG	ATA	TTCC	4068
Qy	2915	Phe	His	Cys	His	Glu	-----	Gly	Tyr	Leu	Leu	His	Gly	Ala	Pro	Lys	2928			
Db	4069	TAC	ACA	TGT	ATG	ACCC	CCAC	CCAG	CAG	AGG	ATG	ACCT	TTC	AACT	TTC	ATTG	GGG	AGAG	CAC	4128
Qy	2929	Leu	Thr	Cys	Gln	Ser	Asp	-----	Gly	Asn	Trp	Asp	Ala	Glu	Leu	Pro	Lys	2944		
Db	4129	ATC	CG	TGCT	GC	ACA	AGT	GC	ACCT	CAT	GGG	AA	TGG	GGT	TGG	AGC	AGC	CCCT	CGCT	4188
Qy	2945	-----	-----	Lys	Pro	Val	Asn	Cys	Gly	Pro	Pro	Glu	Asp	Leu	Ala	His	Gly	Phe	Pro	2960
Db	4189	GAA	CTT	CTT	CTG	TGCT	GGT	CAC	TGT	CTA	ATAA	CCCC	CAG	CAG	CAG	TGTT	CCAT	TGTT	CGC	4248
Qy	2961	-----	-----	Asn	Gly	Phe	Ser	Phe	Leu	His	Gly	His	Leu	Gln	Tyr	Gln	Cys	Phe	2976	
Db	4249	ACG	ATC	CCA	TAA	TAA	TAG	ACT	TTT	GAG	TTT	CCG	AGT	CGG	ACA	TCT	TTGA	ATT	ATG	4308
Qy	2977	Pro	Gly	Tyr	-----	Lys	Leu	His	Gly	Asn	Ser	Ser	Arg	Arg	Cys	Leu	Ser	Asn	Gly	2994
Db	4309	CTG	GGT	ATT	TTT	TGG	AAA	ATA	TGTT	CTCT	ATCTCC	-----	TGC	TAG	AAA	ACT	TGG	TCT	4359	
Qy	2995	Trp	Ser	Gly	Ser	Ser	Pro	Ser	Cys	Leu	Pro	Cys	Arg	Cys	Ser	Thr	Pro	Val	Leu	3014
Db	4360	TGG	T	CAAG	TGTT	GAAG	CAAC	ACT	TAG	ACG	AAAA	ATCAT	TGT	GGAC	CTCC	CAC	CAG	AACTTC	4419	
Qy	3015	-----	Gly	Thr	Val	Asn	-----	Gly	Thr	Asp	Phe	Asp	Cys	Gly						

Db	5077	GACCTCAGAGGGCGTCTCGTCTCTGCACTGCACGCCCGAGGAGACTGGAGCCT	5130
Qy	3232	SerAspGluSerCysSerProValSerCysGlyLys-----ProGluSerProGluHis	3249
Db	5131	GAAGCCCTAGATGTACAGTGAATACTCTGTGATGACTTCTCTGGGCCCACTCCCT	5187
Qy	3250	GlyPheVal-----ValGlySerLysIysThrPheGluSerThr	3262
Db	5188	GGCCGGTGTCTACTTCACCTAATCTCCAGCTTGGGGCAAGGTGTCCTTT	5238
Qy	3263	IleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys	3281
Db	5239	-----GTTTGGCATGAGGGTTCGANTAAAGGCAGGCTCTGCTAGTCATTGTGTC	5289
Qy	3282	-----GlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCys	3299
Db	5290	TTGGCTGGAAATGAAGCCCTTTTGGAAATAGCAGTGTTCAGTGTGTGAACAAATCTTTGT	5349
Qy	3300	GluThrProLeuGluPheLeuAsnGlyLys-----AlaAspIleGlu	3313
Db	5350	CCAAATCTCCAGCTATCTTAATGGGAGACACACAGGAATCCCTTTGGAGATATT	5406
Qy	3314	AsnArgThrThrGlyProAsnValValTyrSerCys-----AsnArgGly	3328
Db	5407	-----CCCTATGGAAAGAAATATCTTAGCATGGACACACCCACGACAGGGATG	5460
Qy	3329	---TyrSerLeuGluGlyProSerGluAlaHisCysThrGlu-----AsnGly	3343
Db	5461	ACCTTCAACCTCATTTGGGAGAGCTCCATCGCTGCACCAAGTGACCCCTCAAGGGAAATGG	5520
Qy	3344	ThrTrpSerHisProValProLeuCysLys-----ProAsnProCysProValPro	3360
Db	5521	GTTTGGAGCAGCCTGCCCTCCCTGTGAACCTTCTGTCTCTGCTGCTGCCACATCCA	5580
Qy	3361	PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal	3380
Db	5581	CCCAAGATC-----CAAAACGGGCATTACATTGGAGACACGTA	5619
Qy	3381	SerIle-----LysCysArgGluGlyPheLeuLeuGln	3391
Db	5620	TCTCTATATCTCTGGGATGACAATCAGCTACACTTGTGACCCCGGCTACCTGTTAGTG	5679
Qy	3392	GlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCys	3411
Db	5680	GGAAAGGGCTCATTTCTGTACAGACCGGGAACTCTGGAGCCAATTTGGATCATTTGCT	5739
Qy	3412	GluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHis	3431
Db	5740	AAAGAAGTAAATTTAGCTTCCCACTGTTTATG---AATGGAATCTCGAAGGAGTTAGAA	5796
Qy	3432	-----TyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMet	3447
Db	5797	ATGAAAAAGTATATCATATGTAGATATTATGTGACTTTGAAGTGTGAAGATGGTATACT	5856
Qy	3448	LeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProIle	3467
Db	5857	CTGGAGGCAGTCCCTGGAGCCAGTCCAGGCGGATGACAGATGG---GACCCTCTCTCTG	5913
Qy	3468	CysArgAlaValCysArg	3473
Db	5914	GCCAAATGTACCTCTCGT	5931

RESULT 8
5256642-3
: Patent No. 5256642
: APPLICANT: PEARSON, DOUGLAS T.; KLINKSTEIN, LLOYD B.; WONG,
: WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; P, STEPHEN
: H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
: TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
: RECEPTOR 1 (CR1) AND A THROMBOCYTIC AGENT, AND THE METHODS OF
: THEREOF
: NUMBER OF SEQUENCES: 30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO: 3
; LENGTH: 5420
5256642-3

Alignment Scores:
Pred. No.: 3,07e-122 Length: 5420
Score: 1604.50 Matches: 475
Percent Similarity: 37.95% Conservative: 191
Best Local Similarity: 27.07% Mismatches: 646
Query Match: 8.03% Indels: 443
DB: 6 Gaps: 74

US-09-977-053-4 (1-3571) x 5256642-3 (1-5420)
Qy 1863 PheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAsp 1882
Db 33 TTTCCTATGGGACATCTTTAAAGTAGCAATGCCGCTCTGAGTAC---TACGGGAGGCCA 89
Qy 1883 LysGluSerSerCysLeuAlaAsnSerSerTyrPheHisSerProProValCysGluPro 1902
Db 90 TTCTCTATCAGATGTAGATTAACCTGTGCTGCTGCTCAAGTCCCAAGAGATGCTGTAAACGT 149
Qy 1903 ValLysCysSerSerProGluAlaAsnLeuAsnGlyLys---TyrLeuSerGlyLeu 1921
Db 150 AAATCATGTAAACCTCTCCAGATCCAGTGAATGGCATGGTGCATGTCATGATCAGACATC 209
Qy 1922 ThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer 1941
Db 210 CAGGTGGTGCAGATCAATCAATCTTCTGTATACAGGCGCCGACTCATTTGGTCACTCA 269
Qy 1942 IleIleGluCysThrAlaSerGly-----IleTyrAspArgAlaProAlaCys 1958
Db 270 TCTGCTGAATGTATCTCTCGGGCATGCTGCCCATTTGGAGCAGCAGCGCCCAATTGT 329
Qy 1959 HisLeuValPheCysGlyGluProProAlaIleLys-----AspAlaValIleThrGly 1976
Db 330 CAACGAATTCCTGTGGGTACCCGCCACCATCGCCAAATGGAGATTTCATTAGCACCAAC 389
Qy 1977 ---AsnAsnPheThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeu 1995
Db 390 AGAGAGAAATTTTCACTATGGATCGTGGTGACCTACCGCTGCAATCTCGGA-----AGC 443
Qy 1996 AlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGln 2015
Db 444 GGAGGGAGAAAGGTGTTGAGCTTGTGGGT-----473
Qy 2016 CysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerProGluThrAla 2035
Db 474 ---GAGCGCTCCATA-----485
Qy 2036 HisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAsp 2055
Db 486 -----TACTGCACC-----AGCAT 500
Qy 2056 AsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTyrValProProGluGlyGlnAspMet 2075
Db 501 GAGATCAAGTG-----GGCATCTGGAGGCGCCG-----GCC 533
Qy 2076 ProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGlu 2095
Db 534 CCTCAGTCATTATACCTAAACAATGACAGCCCTCCAAATGTGGAAATATGGAATATTGTA 593
Qy 2096 SerValSerLysAlaLysPheAlaAlaGlySerValSerPheLysCysMetGluGly 2115

```

```

Db 594 TCTGACACAGAGCTTATTTTCTTAATGAAGTTGTGGAGTTTAGGTGTACGCTGGC 653
Qy 2116 PheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTyrAsnProSer 2135
Db 654 TTTGTGATGAAGAGACCCCGCGTGTGAAGTCCAGGCCCTGAACAAATGGAGCGGAG 713
Qy 2136 ProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGly 2155
Db 714 CTACCAAGC-----TGCTCCAGGGTA---TGTCAGCCACCTCCAGATGTCTCGCATGCT 764
Qy 2156 -----TyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
Db 765 GAGCGTACCCAAAGGAGACAGCAACTTTTCCCGGGCAGGAAGTGTTCACAGCTGT 824
Qy 2173 AsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThrGlyGlnTyr 2192
Db 825 GAGCCCGGCTATGACCTCAGAGGGGCTGCTGTATGCGTGCACACCCAGGGAGACTGG 884
Qy 2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
Db 885 AGCCCTGCAGCCCCCACATGTGAAGTGAATCTCTGTATGAC-----926
Qy 2213 GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn 2232
Db 927 ---TTCAAG-----GGCCAACTTCTTAATGGCCGTGTCTATTT-----962
Qy 2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTyrHis 2252
Db 963 -----CCAGTA-----968
Qy 2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIleGlnAsn 2272
Db 968 -----968
Qy 2273 GlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPheCysAsnGlu 2292
Db 969 -----AATCTCCAGCTTGAGCAAAAGTGGATTTTGTGTGATGAA 1010
Qy 2293 GlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyrAsnLys 2312
Db 1011 GGAATTCATTAATAAAGGCAGCTCTGCTAGTTATGT-----1046
Qy 2313 LysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGlnAsnGln 2332
Db 1046 -----1046
Qy 2333 LeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGluGly 2352
Db 1046 -----1046
Qy 2353 HisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyrPheAspSer 2372
Db 1047 ---GTCTTGGCTGGAATGGAAGCCTT-----TGAATAGCAGT 1082
Qy 2373 PheProValCysLysIleValLeuCysThrProProProLeuIleSerPheGly-----2390
Db 1083 GTTCAGGTGTGTAATAAATCTTTTGTCCAACTCTCCAGTATTCTTAATGGGAGACAC 1142
Qy 2391 ValProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCys-----2408
Db 1143 ACAGGAAACCTCTCGAAGTCTTTCCCTTTGAAAGCAGTAATAATTACATGCGACCCC 1202
Qy 2409 -----ValGlyCysPhePheLeuArgGlyAsnSerThrThrLeuCys-----2422
Db 1203 CACCCAGACAGAGGACGAGCTTCCAGCTCATTTGGAGAGACACCATCCGCTGCACAAGT 1262
Qy 2423 -----GlnProAspGlyThrTyrSerProLeuProGluCys---ValProValGlu 2439
Db 1263 GACCTCAAGGAATGGGTGTGGAGACGCTCCCTCGCTGTGGATTCGGGTGCAC 1322
Qy 2440 CysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyr---2458

```

1323 TGTCAAGCCCGCAGATCATTTCTGTTTCCCAAGTTGAABAAACCCAAACCAATGCATCTGAC 1382
Db
2459 -----LeuSerThrAlaLeu-----TyrThrCysIysProGlyPheGluLeuValGlyAsn 2475
Qy
1383 TTTCCCAATGGGACATCTTTAAAGTAGCAATCGCGTCT-----GAGTACTACGGGAGG 1436
Db
2476 ThrThrThrLeu---CysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLys 2494
Qy
1437 CCAATCTCTATCACAATGCTAGATACTGCTGGTCAAGTCCCAAGATGCTGTAAA 1496
Db
2495 AlaIleGluCysLeuLysProLysGluLeuLeuAsnGlyLysPheSer---TyrThrAsp 2513
Qy
1497 CGTAATCATGTAATAACTCTCCAGATCCAGTGAATGGCATGGTGTGATGTGATCAGAC 1556
Db
2514 LeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgGluGlyPro 2533
Qy
1557 ATCCAGGTTGGATCCAGAATCAACTATCTTGACTACAGGGCCACCGACTCATTTGGTCC 1616
Db
2534 SerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAla 2553
Qy
1617 TCATCTGCTCAATGTATCTCTCAGGC-----ATACT 1649
Db
2554 IleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSer 2573
Qy
1650 GCCCAT----- 1655
Db
2574 TyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMet 2593
Qy
1655 ----- 1655
Db
2594 GlnThrCysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIleAspCys 2613
Qy
1656 -----TGAGCACCAGAGCCGCCAATTTGTCAACGAATCTCTGT 1694
Db
2614 GlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspGlnGlyTyr 2633
Qy
1695 GGGCTACCCCAACCATGGCCCAATGGAGATTTCATT----- 1730
Db
2634 PheGluGlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHis 2653
Qy
1730 ----- 1730
Db
2654 LeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSer 2673
Qy
1731 -----AGCACCAACAGAGAG 1745
Db
2674 AsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGly----- 2688
Qy
1746 AATTTTCATATGATCAGTGGTGACCTACGCTGCAATCTTGGAGCAGAGGGAGAAAG 1805
Db
2689 ---TyrGluLeuLeuGlyAsnProValLeuIleCysGlnGluAsp-----Gly 2703
Qy
1806 GTGTTTGAGCTTGCGGTGAGCCCTCCATATCTGCACCGCAATGATCATCAAGTGGGC 1865
Db
2704 ThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaPro 2723
Qy
1866 ATCTGGAGGGCCCCCTTCAGTGCATTATACCTAACAAATGACAGCCCTCCAAATGTG 1925
Db
2724 GluAsnGlyPheLeuArgPheThrGluThrSerMetGlySer-----AlaValGln 2740
Qy
1926 GAAATGGATATGGTATCTGCACACAGAAAGCTTATTTCTTAAATGAATGTGTGGAG 1985
Db
2741 TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
Qy
1986 TTTAGGTGTGAGCTGCTGTTGTGATCAAGAGAGACCCCGCGTGTGAGAGTCCAGGCCCTG 2045
Db
2761 ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
Qy
2046 AACAAATGGAGGCCAGATTACCAAGCTGCTCCAGGGGTG---TGTACGCCCTCCAGAA 2102
Db
2781 ValMetAsnGlySerIleLysGlySer-----AsnTyrThrTyrLeuSerThrLeu 2797
Qy
2103 ATCTGCATGTTGAGCATATACCCAGAGCCATCAGGACAACTTTTTCACCTGGGCGAGAGTG 2162
Db

2798 TyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgThrCysGlnAsp 2817
Qy
2163 TTTCTACAGCTGTGAGCCCTGGCTATGACTCAGAGGGCTGGTCTCTGCTACTGACACCC 2222
Db
2818 AspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProPro 2837
Qy
2223 CGAGGACAGTGGAGCCCTCGAAGCCCGAGATGTGCAGTGAATACTCTGTGAT----- 2273
Db
2838 ValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGln----- 2852
Qy
2274 ---GACTTCTTGGTCAACTCCCTCATGGCGGTGTGCTATTTCACATTAATCTCCAGCTT 2330
Db
2853 ---LysGluIleGluTyrThrCysAsnGlyGlyPheLeuLeuGluGlyValaArgSerArg 2871
Qy
2331 GGGCAAGGTGCTCTTGTCTGTGATGAGGGTTCGCTTAAAGGGCGGTTCCTGTAGT 2390
Db
2872 ValCysLeuAlaAsnGlySer-----TrpSerGlyAlaThrProAspCysValPro 2888
Qy
2391 CATTTGTCTTGGTGGATCAGAGAGCCTTTTGAATAACAGTGTTCCTGTGTGTAACAT 2450
Db
2889 ValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeu----- 2905
Qy
2451 ATCTTTTGTCCAAATCTCTCAGCTATCTTAAATGGGAGACACACAGCACTCCCTCTGGA 2510
Db
2906 AspTyrGlyPheMetLysGluValThrPheHisCysHisGlu----- 2919
Qy
2511 GATATTCCTTGGAAAGAAATATCTTACATGTGTGACCCCAACCCAGACAGAGGGATG 2570
Db
2920 GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAsp-----Gly 2935
Qy
2571 ACCTTCAACTCATTTGGGAGAGCACCATCGCTGCACAGTGTACCTCATGGGAATGGG 2630
Db
2936 AsnTrpAspAlaGluIleProLeuCys-----LysProValAsnCysGlyPro 2951
Qy
2631 GTTTGGAGCAGCCCTGCTCGCTGTGAACCTTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 2690
Db
2952 ProGluAspLeuAlaHisGlyPhePro-----AsnGlyPheSerPheIleHis 2967
Qy
2691 CCGAGCAGGTTTCCATTTGGCAGTCTTACGATCCCAATTAATGACTTTGAGTTCCAGTC 2750
Db
2968 GlyGlyHisIleGlnTyrGlnCysPheProGlyTyr-----LysLeuHisGlyAsnSer 2985
Qy
2751 GGGACATCTTTGAATTAATGAATCGCTCGGTATTTTGGGAAAAATCTTCTATCTCC 2810
Db
2986 SerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCys 3005
Qy
2811 -----TGCCTAGAAAACTTGGTCTGTCAAGTGTGAAGACAACTGTAGACGAAAA 2861
Db
3006 ArgCysSerThrProValIleGluTyr---GlyThrValAsn---GlyThrAspPheAsp 3023
Qy
2862 TCATGTGGACTCCACAGAACCTTTCATAGGAATGTGTGCATATAACACAGATACACAG 2921
Db
3024 CysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGlu 3043
Qy
2922 TTTGGATCAACAGTTAATTAATCTTGTAAAGGGTTTCGACTCATTTGGTTCCTCATCT 2981
Db
3044 IleThrCysGluAlaAspGlyGln-----TrpSerSerGlyPheProHisCysGlu 3060
Qy
2982 ACTACTGTCTGTCTCAGGCAATAATGTGCATGGGATAAGAGGCACTTATTGTGAG 3041
Db
3061 HisThrSerCysGlySerLeuProMetIleProAsnAla---PheIleSer-----Glu 3077
Qy
3042 ATCATATCTTGTGAGCCACTCCAAACCATATCCAAATGGAGACTTCTACAGCAACATAGA 3101
Db
3078 ThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGly----- 3093
Qy
3102 ACATCTTTTCAATGGAACGGTGGTAACTTACCAGTGCCACACTGGACAGATGGAGAA 3161
Db
3094 -----TyrValIleGlnGlySerSerAspLeuIleCysThrGluLys----- 3107
Qy
3162 CAGCTGTTTGAGCTTGTGGGAGAACGGTCAATATATTGACACAGCAAGATGATCAAGTT 3221
Db


```

QY 3108 GlyValTyrSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSer 3127
Db 3222 GGTGTTGGAGCAGCCCTCCCTCGGTGTAATTTCTACTAAATGACACAGCTCCAGAA 3281
QY 3128 ValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSer-----GluVal 3144
Db 3282 GTTGAAATGCAATTTAGAGTACCAGAAACAGAGAGTTCTTTTCTCCTACTGAGATCATC 3341
QY 3145 LysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrPheThrCysGln 3164
Db 3342 AGATTAGATGTCAGCCCGGGTTTTCATG---GTAGGGTCCACACACTGTGCAGTGCAG 3398
QY 3165 LysAspGlyArgTyrPheProGluArgIleSerCysSerProLysCysProLeuPro 3184
Db 3399 ACCAATGGCAGATGGGGGCCCAACACTGCTCCAGGTGTGTGAGCGGCTCCCA 3458
QY 3195 GluAsnIle-----ThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGln 3202
Db 3459 GAAATCTGCATGTGTGAGCATACCTAAGCCATCAGGACACTTTTTCACCTGGGCGAGAA 3518
QY 3203 ValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCysGln 3222
Db 3519 GTGTTCTACAGCTGTGAGCCAGCTATGACCTCAGAGGGGCTGCTCTGTGCACTGCAG 3578
QY 3223 LeuAspGlyThrTyrPheProGluSerAspGluSerCysSerProValSerCysGly 3242
Db 3579 CCCAGGGAGACTGGAGCCCT-----GAAGCCCTAGATGTACAGTGAATCATCTGTGAT 3632
QY 3243 Lys-----ProGluSerProGluHisGlyPheVal-----Val 3253
Db 3633 GACTTCTCGGGCCAACTCCCT---CATGGCCGTGTGCTACTTCCACTTAATCTCCAGCTT 3689
QY 3254 GlySerLysThrPheGluSerThrIleLeuTyrGlnCysGluProGlyTyrGluLeu 3273
Db 3690 GGGCAAGAGTGTCTTT-----GTTTGGCATGAAGGGTTCGGAATTA 3731
QY 3274 GluGlyAsnArgGluArgValCys-----GlnGluAsnArgGlnTyrSerGlyGly 3290
Db 3732 AAAGCAGCTGTGCTAGTCAATGTTGTCTTGGCTGGAATGAAGCCCTTTGGGAATGACGT 3791
QY 3291 ValAlaIleCysLeuGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLys--- 3309
Db 3792 GTTCAGTGTGTGAACAATCTTTGTCCAAATCTCCAGCTATCTCTTAATGGGAGACAC 3851
QY 3310 -----AlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSer 3324
Db 3852 ACAGGAACCTCCCTTTGGAGATATT-----CCCTATGGAAGAAATAATCTTACGCA 3902
QY 3325 Cys-----AsnArgGly-----TyrSerLeuGluGlyProSerGluAlaHis 3338
Db 3903 TGCAGACCCACCCAGACAGAGGGATGACCTTCAACCTCATTTGGGGAGAGCTCCCATCCGC 3962
QY 3339 CysThrGlu-----AsnGlyThrTyrPheHisProValProLeuCysLys--- 3353
Db 3963 TGCACAGTGACCCCTCAAGGGATGGGTTTGGAGCGCCCTGCCCTCGCTGTGNACTT 4022
QY 3354 -----ProAsnProCysProValProPheValIleProGluAsnAlaLeuLeuSerGlu 3371
Db 4023 TCTGTTCTGCTGCTGCCCCACATCCACCCCAAGATC-----CAA 4061
QY 3372 LysGluPheTyrValAspGlnAsnValSerIle----- 3382
Db 4062 AACGGGCATTATCATGGAGACAGCATCTCTATATCTTCTCGGGATGACATCAGCTAC 4121
QY 3383 LysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleThrCysAsnProAspGlu 3402
Db 4122 ACTTGTGACCCCGGTACTCTGTAGTGGGAAGGGCTTCATTTCTGTACAGACCCAGGA 4181
QY 3403 ThrTyrThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisVal 3422
Db 4182 ATCTGGAGCCAAATGGATCATTTATTCAAAGAGTAATTTAGCTTCCCACTGTTTATG 4241
QY 3423 GluAsnAlaIleAlaArgGlyValHis-----TyrGlnTyrGlyAspMetIle 3438

```

```

Db 4242 ---AATGGAATCTCGAAGAGTTAGAAATGAAAAAGATATATCATCTATGGAGATTATG 4298
QY 3439 ThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGlu 3458
Db 4299 ACTTTGAAGTGTGAAGATGGGTATATCTCTGGAAGGAGTCCCTGGAGCCAGTCCAGGCG 4358
QY 3459 AsnGlyThrTyrThrSerProProIleCysArgAlaValCysArg 3473
Db 4359 GATGACAGATGG---GACCCCTCTCTGGCCAAATGATACCTCTCGT 4400

RESULT 9
5472939-3
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; COMCINO, MICHAEL F.; IP, STEPHEN
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138,825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588,128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO: 3
; LENGTH: 5420
5472939-3

Alignment Scores:
Score: 3,07e-122 Length: 5420
Pred. No.: 1604.50 Matches: 475
Percent Similarity: 37.95% Conservative: 191
Best Local Similarity: 27.07% Mismatches: 646
Query Match: 8.03% Indels: 443
DB: 6 Gaps: 74

US-09-977-053-4 (1-3571) x 5472939-3 (1-5420)
QY 1863 PheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAsp 1882
Db 33 TTTCCCATTTGGGACATCTTTAAAGTACGAATGCCGTCTGAGTAC---TACGGGAGGCCA 89
QY 1883 LysGluSerSerCysLeuAlaAsnSerSerTyrPheHisSerProProValCysGluPro 1902
Db 90 TTTCTCTATCATCATGTCTAGATAACCTGCTGTGTCAGATGCCCAAGATGTTCTGTAAACGT 149
QY 1903 ValLysCysSerSerProGluAsnIleAsnAsnGlyLys---TyrIleLeuSerGlyLeu 1921
Db 150 AAATCATGTAAACCTCTCCAGATCCAGTGAATGGCATGTGTGATGTGATCACAGACATC 209
QY 1922 ThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer 1941
Db 210 CAGTTGGATCCGAATCAACTATTCTTGTACTACAGGGCAGCCGACTCATTTGCTCACTCA 269
QY 1942 IleIleGluCysThrAlaSerGly-----IleTyrAspArgAlaProProAlaCys 1958
Db 270 TCTGCTGAATGTATCTCTCGGGCAATGTGCCCATTTGGAGCAGCAGCGCCCAATTGT 329
QY 1959 HisLeuValPheCysGlyGluProProAlaIleLys-----AspAlaValIleThrGly 1976
Db 330 CAAAGCAATTCCTTGTGGGCTTACCCCAACCATGCCCCAATGGAGATTCATTAGCACCAAC 389
QY 1977 ---AsnAsnPheThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeu 1995
Db 390 AGAGAGAAATTTTCACTATGATCAGTGGTGCACCTACCGCTGCAATCTCTGGA-----AGC 443

```

QY 1996 AlaGlyLeuAspThrIleGluCysLeuAlaAspGlyIleTyrSerArgSerAspGlnGln 2015
DB 444 GGAGGGAGAAAGTGTGGCTGGTGGT----- 473
QY 2016 CysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerProGluThrAla 2035
DB 474 -----GAGCCCTCCATA----- 485
QY 2036 HisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAsp 2055
DB 486 -----TACTGCACC-----AGCAAT 500
QY 2056 AsnSerGlnLeuLeuCysAsnAlaGlnGlyIleTyrValProProGluGlyGlnAspMet 2075
DB 501 GAGCATCAAGTG-----GGCATCTGGAGCGGCCG-----GCC 533
QY 2076 ProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGlu 2095
DB 534 CCTCAGTGCATTATACCTAACAAATGACCGCTCCAAATGTGGAATGGAATATTGGTA 593
QY 2096 SerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGly 2115
DB 594 TCTGACACACAGAGCTTATTCTCTTAATGAAGTTGTGGAGTTTAGGTGTCAAGCTGGC 653
QY 2116 PheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTyrAsnProSer 2135
DB 654 TTTGTCTGAAGAGACCCCGCTGTGAAGTGCAGGCCCTGGAACAAATGGGAGCGGAG 713
QY 2136 ProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGly 2155
DB 714 CTACCAAGC-----TGCTCCAGGTA---TGTCAGCCACTTCAGATGCTCTGCATGCT 764
QY 2156 -----TyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
DB 765 GAGCGTACCCAAAGGACAGGACAACTTTTACCCTCCGCGGAGGAGTGTCTACAGCTGT 824
QY 2173 AsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThrGlyGlnTyr 2192
DB 825 GAGCCCGGCTGATGACCTCAGAGGGGTGCTCTATGCGCTGCACACCCAGGAGACTGG 884
QY 2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
DB 885 AGCCCTGCACCCCACTGTGAAGTGAATCTGTGTATGAC----- 926
QY 2213 GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn 2232
DB 927 ---TTCATG-----GGCCAACTTCTTAATGGCGCTGTGTATTT----- 962
QY 2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTyrHis 2252
DB 963 -----CNGTA----- 968
QY 2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIleGlnAsn 2272
DB 968 ----- 968
QY 2273 GlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
DB 969 -----AATCTCCAGCTGGAGCAAAAGTGGATTTTGTGTGATGA 1010
QY 2293 GlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyrAsnLys 2312
DB 1011 GGATTTCAATTAAGAGGACGCTCTGTGTATTGT----- 1046
QY 2313 LysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsnGln 2332
DB 1046 ----- 1046
QY 2333 LeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGluGly 2352
DB 1046 ----- 1046

QY 2353 HisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyrAsnAspSer 2372
DB 1047 ---GYCTTGGCTGGATGGAAAGCCCT-----TGGAAATAGCAGT 1082
QY 2373 PheProValCysLysIleValLeuCysThrProProProLeuIleSerPheGly----- 2390
DB 1083 GTTCCAGTGTGTGAACAAATCTTTTGTCCAAAGTCTCCAGTATTCTTAATGGGAGACAC 1142
QY 2391 ValProIleProSerSerAlaLeuHisPheGlySerThrValIleTyrSerCys----- 2408
DB 1143 ACAGGAAAACCTCTGGAAGTCTTTCCCTTTGGAAAGACAGTAAATTACATGCGACCCC 1202
QY 2409 -----ValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCys----- 2422
DB 1203 CACCACACAGAGGAGCGAGCTTTCGACCTCATTTGGAGAGAGCACCATCCGCTGCACAAGT 1262
QY 2423 -----GlnProAspGlyThrTyrSerSerProLeuProGluCys-----ValProValGlu 2439
DB 1263 GACCCTCAAGGGAATGGGGTTTGGAGCAGCCCTGCCCTGTGCTGGAATTTCTGGGTCCAC 1322
QY 2440 CysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyr--- 2458
DB 1323 TGTCAAGCCCAATCATTTTCTGTTCGCAAGTTGAAACCCCAACCAATGCAATCTGCAC 1382
QY 2459 -----LeuSerThrAlaLeu---TyrThrCysLysProGlyPheGlyLeuValGlyAsn 2475
DB 1383 TTTCCCAATGGGACATCTTTAAAGTACGAATCGCTCT-----GAGTACTACGGGAGG 1436
QY 2476 ThrThrThrLeu---CysGlyGluAsnGlyHisTyrLeuGlyLysProThrCysLys 2494
DB 1437 CCATTCTCTATCATGTCTAGATAACCTGTGTCTGTCAAGTCCCAAGATGTCTGTAAA 1496
QY 2495 AlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSer---TyrThrAsp 2513
DB 1497 CGTAATCATGTAAACTCTCCAGATCCAGTGAATGGCATGGTCATGTCATCACACAC 1556
QY 2514 LeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyPro 2533
DB 1557 ATCCAGTTCGATCCAGAAATCAACTATTCTGTACTACAGGCGCAGCTCATTTGGTCCAC 1616
QY 2534 SerAlaLeuThrCysLeuGluThrGlyAspTyrPheValAspAlaProSerCysAsnAla 2553
DB 1617 TCATCTGCTGAATGTATCTCTCAGGC-----AATACT 1649
QY 2554 IleHisCysAspSerProGlnProIleLeuAsnGlyPheValGluGlyAlaAspTyrSer 2573
DB 1650 GCCCAT----- 1655
QY 2574 TyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMet 2593
DB 1655 ----- 1655
QY 2594 GlnThrCysGluGluSerGlyTyrSerSerIleProThrCysMetProIleAspCys 2613
DB 1656 -----TGGAGCACGAAGCCGCAATTTGTCAACGAATTTCTTGT 1694
QY 2614 GlyLeuProProIleAlaAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyr 2633
DB 1695 GGGCTACCCCAACCACTCCGCAATGGAGATTTCAAT----- 1730
QY 2634 PheGluGlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHis 2653
DB 1730 ----- 1730
QY 2654 LeuGlyAlaValAlaLysThrTyrGluAsnThrLysGluSerProAlaThrHisSerSer 2673
DB 1731 -----AGCACCAACACAGAGAG 1745
QY 2674 AsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGly----- 2688
DB 1746 AATTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
QY 2689 ---TyrGluLeuLeuGlyAsnProValLeuIleCysGlnGluAsp-----Gly 2703

Db 1806 GTGTTTGAGCTTGTGGGTGAGCCCTCCATATCTACTGACACGACATGACATGAGTGGCC 1865
QY ThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaPro 2723
Db 1866 ATCTGGAGCGGCCCGCCCTCAGTGCATTAATACCAATATGACAGCCCTCCAAATGTG 1925
QY GluAsnGlyPheLeuArgPheThrGluThrSerMetGlySer-----AlaValGln 2740
Db 1926 GAAATATGGAATATTTGATCTGACACAGAGACTTATTTTCTTAAATGAAGTTGTGGAG 1985
QY TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
Db 1986 TTAGGTGTGACCTTGCTTGTCTGATGAAGAGACCCCGCGTGTGAAGTGCACAGCCCTG 2045
QY ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
Db 2046 AACAAATGGAGCCAGAGTGTACCAAGCTGTCCAGGGTG---TGTCAGCGCCCTCCAGAA 2102
QY ValMetAsnGlySerIleLysGlySer-----AsnTyrThrTyrLeuSerThrLeu 2797
Db 2103 ATCTGTGATGTGAGCATACCCCAAGCCATCAGGACAACTTTTTCACCTGGGCGAGGATG 2162
QY TyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAsp 2817
Db 2163 TTCTACAGCTGTGAGCTTGCTGTGATGAGGGTTTCGCTTAAGGGCAGTTCCGTAGT 2390
QY AspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProPro 2837
Db 2223 CGAGGAGACTGAGCCCTGGAAGCCCGAGATGTGCAGTGAATCTCTGTGAT 2273
QY ValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPheGln----- 2852
Db 2274 ---GACTTCTTGGGTCAACTCCCTCATGCGCCGTGCTATTTCCACTTAATCTCCAGCTT 2330
QY ---LysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArg 2871
Db 2331 GGGCAAGAGTGTCTTGTCTGTGATGAGGGTTTCGCTTAAGGGCAGTTCCGTAGT 2390
QY ValCysLeuAlaAsnGlySer-----TrpSerGlyAlaThrProAspCysValPro 2888
Db 2391 CATTTGTCTTGGTGAATGAGAGCCCTTGGAAATACAGTGTCTCTGTGTGTAACAT 2450
QY ValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeu----- 2905
Db 2451 ATCTTTTGTCCAAATCTCTCAGTATCTCTTAATGGGAGACACACAGGAACCTCCCTCTGGA 2510
QY AspTyrGlyPheMetLysGluValThrPheHisCysHisGlu----- 2919
Db 2511 GATATTCCTTGGAAAGAAATATCTACATGTGACCCCAACCCACAGAGGGATG 2570
QY GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAsp-----Gly 2935
Db 2571 ACCTTCAACCTCATTTGGGAGAGACCAATCCGCTGCACAAATGACCCCTCATGGGAATGG 2630
QY AsnTrpAspAlaGluIleProLeuCys-----LysProValAsnCysGlyPro 2951
Db 2631 GTTTGGAGAGCCCTGCCCCCTCGCTGTGAACCTTTCTGTGCTGTGCTCAGTGAACCC 2690
QY ProGluAspLeuAlaHisGlyPhePro-----AsnGlyPheSerPheIleHis 2967
Db 2691 CCAGAGCAGTTTCCATTTCCAGTCTACATCCCAATTAATGACTTTTTCAGTTCCAGTC 2750
QY GlyGlyHisIleGlnTyrGlnCysPheProGlyTyr-----LysLeuHisGlyAsnSer 2985
Db 2751 GGGACATCTTTGAATATGAATGATGAGTGTGATTTTGGGAAATGTTCTCTATCTCC 2810
QY SerArgArgCysSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCys 3005
Db 2811 ---TGCTAGAAAATCTGGTCTGGTCAAGTGTGAGACAACTGTAGACGAA 2861
QY ArgCysSerThrProValIleGluTyr---GlyThrValAsn---GlyThrAspPheAsp 3023

Db 2862 TCATGTGGAGCTCCACCAGAACCCCTTCAATGGAATGGTGCATATAAACACAGATACACAG 2921
QY CysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGlu 3043
Db 2922 TTTGGATCAACAGTTAAATTAATTTCTTGTATGAAGGGTTTCGACTCATCTTGGTCCCATCT 2981
QY IleThrCysGluAlaAspGlyGln-----TrpSerSerGlyPhePheProHisCysGlu 3060
Db 2982 ACTACTTGTCTCGTCTCAGCAATATGTCATATGGGATAGAGGACCACTTATTTGTGAG 3041
QY HisThrSerCysGlySerLeuProMetIleProAsnAla---PheIleSer-----Glu 3077
Db 3042 ATCATATCTTGTGAGCCACCTCCACCAATATCCAAATGGAGACTTCTTACAGCAACAATAGA 3101
QY ThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSerGly----- 3093
Db 3102 ACATCTTTTCACAATGGAACGGTGTAACTTACCACTGTCACACTGGACCCAGATGAGAA 3161
QY ---TyrValIleGlnGlySerAspLeuIleCysThrGluLys----- 3107
Db 3162 CAGCTGTTTGAGCTTGTGGAGAACCGTCAATATATTTGCCAGCAAGATGATCAAGTT 3221
QY GlyValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSer 3127
Db 3222 GGTGTTTGGAGCAGCCCTCCCTCGTGTGATTTCTACTAATAAATGCACAGCTCCAGAA 3281
QY ValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSer-----GluVal 3144
Db 3282 GTTGAATAATGCAATTAGTACAGAAACAGGAGTTTCTTTTCCCTCACTGATGATCATC 3341
QY LysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPheThrCysGln 3164
Db 3342 AGATTAGATGTGAGCCCGGGTTTGTCAAG---GTAGGGTCCCACACTGTGCAGTGCAG 3398
QY LysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeuPro 3184
Db 3399 ACCAATGGCAGATGGGGCCCAAGCTGCCACACTGTCTCCAGGGTGTGTTCAGCGCCCTCCA 3458
QY GluAsnIle-----ThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGln 3202
Db 3459 GAAATCTCTGAGTGTGAGCATACCTTAAGCCATCAGGACAACTTTTTCACCTGGCGGAGAA 3518
QY ValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCysGln 3222
Db 3519 GTGTTCTACAGCTGTGAGCCAGCTATGACTCAGAGGGGCTCGCTCTCTGCATGCAAG 3578
QY LeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSerCysGly 3242
Db 3579 CCCAGGAGAGACTGGAGCCCT-----GAAGCCCTAGATGTACAGTGAATACTCTGTGAT 3632
QY Lys-----ProGluSerProGluHisGlyPheVal-----Val 3253
Db 3633 GACTTCTGGGCCAACTCCCT---CATGGCCGTGTCTACTTCCACTTAATCTCCAGCTT 3689
QY GlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeu 3273
Db 3690 GGGCAAGAGTGTCTTT-----GTTTGGATGAAGGGTTCCGATTA 3731
QY GluGlyAsnArgGluArgValCys-----GlnGluAsnArgGlnTrpSerGlyGly 3290
Db 3732 AAAGGAGGTCTGTAGTCAATGTGTCTTGTGGCTGGAATGAAGCCCTTTTGGATAGCAGT 3791
QY ValAlaIleCysValysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLys--- 3309
Db 3792 GTTCCAGTGTGGAACAAATCTTTTGTCCAAATCTCTCCAGTATCTTATTTGGAGACAC 3851
QY -----AlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSer 3324
Db 3852 ACAGGAATCTCTTTTGGAGATAT-----CCCTATGGAAGAAATATCTTACGCA 3902
QY Cys-----AsnArgGly-----TyrSerLeuGluGlyProSerGluAlaHis 3338
Db 3903 TGGCAGACCCCAACCCAGACAGAGGATGACCTTCACTTCACTTGGGGAGAGCTCCATCCGC 3962

3339 CysThrGlu-----benGlyThrTrpSerHisProValProLeuCysLys--- 3353
3963 TGCACAGTGACCTCAAGGAATGGGTTGGAGCAGCCCTCGCTCGTGAACATT 4022
3354 -----ProAsnProCysProValProPheValIleProGluAsnAlaLeuLeuSerGlu 3371
4023 TCTGTTCCTGCTGCTGCCCATCCACCAAGATC-----CAA 4061
3372 LysGluPheTyrValAspGlnAsnValSerIle----- 3382
4062 AACGGCCATTACATTGGAGCAGCAGTATCTCTATATCTTCTGGGATGACAAATCAGCTAC 4121
3383 LysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGlu 3402
4122 ACTTGTGACCCCGCTACCTGTTAGTGGAAAGGGCTTCATTTTCTGACAGCAGGGA 4181
3403 ThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisVal 3422
4182 ATCTGGAGCAATTGATCATTTATTCGAAGAGTAATTTAGTCTCCACTGTTTATG 4241
3423 GluAsnAlaIleAlaArgValHis-----TyrGlnTyrGlyAspMetIle 3438
4242 ---AATGGAATCTCGAAGCAGTTAGAAATGMAAAAGATATATCACTATGAGATTATGTG 4298
3439 ThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGlu 3458
4299 ACTTTGAAGTGTGAAGATGGGTATATCTCTGGAAGGAGTCCCTGGAGCCAGTCCAGGCG 4358
3459 AsnGlyThrTrpThrSerProProIleCysArgAlaValCysArg 3473
4359 GATGACAGATGG---GACCTCTCTGGCCAAATGTACCTCTCGT 4400

RESULT 10

US-09-023-655-1066
Sequence 1066, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
FILING DATE: US/09/023,655
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1066:
SEQUENCE CHARACTERISTICS:

LENGTH: 3934 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g181939
US-09-023-655-1066

Alignment Scores:
Pred. No.: 1.5e-87 Length: 3934
Score: 1184.00 Matches: 342
Percent Similarity: 35.61% Conservative: 142
Best Local Similarity: 25.17% Mismatches: 379
Query Match: 5.93% Indels: 496
DB: 54 Gaps: 54

US-09-977-053-4 (1-3571) x US-09-023-655-1066 (1-3934)

QY 1688 IleSerCysGlyValProProLeuGluAsnGlyPheHisSerAlaAspPheTyr 1707
Db 130 ATTTCTTGCTCTCTCTCGCTATCTTAATGGCCGATAGT-----TATTAT 180
QY 1708 Ala-----GlySerThrValThrGlnCysAsnAsnGlyTyrTyrIleu 1722
Db 181 TCTACCCCATTCCTGTTGGTACCGTAAAGGTACAGTTGTTTCAGGTACCTTCGCGCTC 240
QY 1723 LeuGlyAspSerArgMetPheCysThrAsp-----AsnGlySerTrpAsnGly 1738
Db 241 ATTTGGAGAAAAAGTCTTATTATGATCACTAAAGACAAAGTGGATGAACCTGGGATAA 300
QY 1739 ValSerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHis 1758
Db 301 CTGCTCTCTAAATGT----- 315
QY 1759 AlaSerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGly 1778
Db 316 -----GAATATTTCAT 327
QY 1779 AspGlyLysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGly 1798
Db 328 AATATTCTTTGCTGCTGAGCCNTA-----GTACAGGAGGATACAAATATAGA 378
QY 1799 HisSerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGly 1818
Db 379 GGCTCTACACCC-----TACAGACATGGTGATCTGTGACATTTGCTGTAAACCAAC 432
QY 1819 TyrGlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeu 1838
Db 433 TCTCCATGAACGMAACAAGTCTGTTGGTGTCAAGCAATAATATGTGGGGCCGACA 492
QY 1839 ---IleProTyrCysLysAlaVal-----SerCysGlyLysProAlaIlePro--- 1853
Db 493 CGACTACCAACCTGTGTAGTATTTTCCCTCTCGAGTGT-----CCAGCACTTCTCTATG 546
QY 1854 -----GluAsn---GlyCysIleGluGluLeuAlaPheThr 1864
Db 547 ATCCACATGGACATCACAAGTGAGATCTTGGTCCATTTGCTCCA----- 594
QY 1865 PheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLysGlu 1884
Db 595 ---GGATTGTCTGACTTACAGCTGTGAATCTGTTACTTGTCTGTTGGAGAAAGATC 651
QY 1885 SerSerCysLeuAlaAsnSerSerTrpSerHisSerProValCysGluProValLys 1904
Db 652 ATTAACATGTTTCTTCGGGAAATGAGGTGCTCTCCCCCCACATGTGAAGAGCAGC 711
QY 1905 CysSerSerProGluAsnIleAsnAsnGlyLys-----TyrIleLeuSer 1919
Db 712 TGTAAATCTCTAGGACGATTTCCCAATGGGAAGTAAAGAGGCTCCCAATTTCTCGGTT 771
QY 1920 GlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGly 1939
Db -----

Db 772 GGTGTA-----ACTGCAAACTTTTCTGTGATGAAGGGTATCGACTGCAAGGC 819
Qy 1940 ProSerIleIleGluCysThrAlaSerGly-----IleTrrAspArgAlaProPro 1956
Db 820 CCACCTTCTAGTCGGTGTGTAATTCGTGACAGGAGGTGCTGGACCAAAATG---CCA 876
Qy 1957 AlaCysHisLeuValPheCysGlyGluProProAlaIleLeuAspAlaValIleThrGly 1976
Db 877 GTATGTGAAGAAATTTTGGCCCATCACCTCCCTTATTTCTCAATGAAGACATATATAGC 936
Qy 1977 Aen-----AsnPheThrPheArgAsnThrValThrTyrThrCys----- 1989
Db 937 AACTCACTAGCAAAATGCTCATATGGAAGCATAGTCACTTGTGACCCGGACCCA 996
Qy 1990 LysGluGly-----TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAsp--- 2006
Db 997 GAGGAAGGAGTGAACCTTCATCTTATTGGAGAGACACTCTCCGTTGTACATGTATAGT 1056
Qy 2007 -----GlyTyrTrpSerArgSerAspGlnCys-----LeuAlaVal 2019
Db 1057 CAGAAAGACTGGGACCTGGAGTGGCCCTGCCACGCTGTGAACCTTCTACTTCTCGGGTT 1116
Qy 2020 SerCysAspGluProProIleVal-----AspHisAlaSerProGluThrAlaHisArg 2037
Db 1117 CAGTGTCCACATCCCGACATCTTAAGAGCGCGAATGATCTGGCAGAAAGATCGATAT 1176
Qy 2038 LeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAenSer 2057
Db 1177 ACCTATAACGACACACTGTGATATTTGCTTGCATGTTTGGCTTCACCTTGAAGGGCAGCAAG 1236
Qy 2058 GlnLeuLeuCysAsnAlaGlnGlyTyrTrpValProProGluGlyGlnAspMetProArg 2077
Db 1237 CAATCCGATGCGATGCGCCAGGACATCGGAGCCATCT----- 1275
Qy 2078 CysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerVal 2097
Db 1276 -----GCACCACTCTGTGAAAG----- 1293
Qy 2098 SerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheVal 2117
Db 1293 ----- 1293
Qy 2118 LeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMet 2137
Db 1293 ----- 1293
Qy 2138 SerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAla 2157
Db 1294 -----GAATGCCAGGCGCCCTCTTAACATCTCAATGGCAAAAG 1332
Qy 2158 SerGlySerAsn-----TyrSerPheGlyAlaMetValAlaTyrSerCysAsnLys 2174
Db 1333 GAAGATAGACACATGTCGCTTTGACCTGGAAACATCTATAAAATATATAGCTGTAAACCT 1392
Qy 2175 GlyPheTyrIleLysGlyGluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSer 2194
Db 1393 GCCTATGCTGCTGGAGAGAAATCCATACAGTGTACTCTGAGGGGTGGACACCC 1452
Qy 2195 ProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsnGlyPhe 2214
Db 1453 CCTGTACCCCAATGCAAAGTGGCAGCGTGT----- 1482
Qy 2215 LeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsnProGly 2234
Db 1483 -----GAAGCTACAGGAAG----- 1497
Qy 2235 TyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHisSerGlu 2254
Db 1497 ----- 1497
Qy 2255 SerProLeuMetCysValProLeuAspCysGlyLysProProIleGlnAsnGlyPhe 2274
Db 1498 ---CAACTCTTCACAAACCC-----CAGCACCAAAATTT 1527

Qy 2275 MetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyr 2294
Db 1528 GTTAGCACCATGAT-----GTCACTCTTCTTGTGTGGAGGGTAC 1566
Qy 2295 GluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSer 2314
Db 1567 AAGTTAAGTGGAGGTGTTTATCAGGAGTGTCAAGGCACATTCCTTGG----- 1614
Qy 2315 AsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuVal 2334
Db 1615 -----TTTATGAGATTCCTCTT----- 1632
Qy 2335 LeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGluGlyHisVal 2354
Db 1632 ----- 1632
Qy 2355 LeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPhePro 2374
Db 1632 ----- 1632
Qy 2375 ValCysLysIleValLeuCysThrProProLeuIleSerPheGlyValProIlePro 2394
Db 1633 ---TGTAAAGAAATCACCTGCCACCCCTCTTATCTACAATGGGCA----- 1680
Qy 2395 SerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGlyPhePheLeu 2414
Db 1681 -----CACACCGGAGTTCCTTA----- 1698
Qy 2415 ArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGlu 2434
Db 1698 ----- 1698
Qy 2435 CysValProValGluCysProGlnProGluGluIleProAsnGlyIleIleAspValGln 2454
Db 1699 -----GAAGATTTTCCATATGCA----- 1716
Qy 2455 GlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGly----- 2469
Db 1717 -----ACCACGGTCACTTACACATGTAACTTTCCCTCTGCTGTCAGTGC----- 1761
Qy 2470 ---PheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsn-----Gly 2484
Db 1762 GAATTCAGCTCATTTGGAGAGAGACCATCCCGTTGTACAAGCAATGATCAAGAAGAGGC 1821
Qy 2485 HisTrpLeuGlyGlyLysProThrCysLys-----AlaIleGluCysLeuLys 2500
Db 1822 ACCTGGAGTGGCCCTGCTCCCTATGTAACTTTCCCTCTGCTGTCAGTGC-----TCA 1878
Qy 2501 ProLysGluIleLeuAsnGly---LysPheSerTyrThrAsp-----LeuHisTyrGly 2517
Db 1879 CATGTCCATTTTGCATATGGATACAGATATCTGGCAAGAGACCCCATATTTCTACAT 1938
Qy 2518 GlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThr 2537
Db 1939 GACACTGTGACATTCAGTGTATAGTGGATTTACTTTGAAGCGCATAGTACATTCGT 1998
Qy 2538 CysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAsp 2557
Db 1999 TGCNAAGCTGATAACACCTGGGATCTCGAAATACCACTTTGTGNAAGAAAGAACATGC----- 2055
Qy 2558 SerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIle 2577
Db 2056 -----CAGCATGTGACAGACAGATCTTCAAGAACTTCCAGCTGGTTCACGTGGAGCTA 2109
Qy 2578 IleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGlu 2597
Db 2110 GTTAATACCTCTCGCAAGATGGTACCACTGTTGACTGGACATGCTTATCAGATGTGTCAA 2169
Qy 2598 -----GluSerGly---TrpSerSerSerIleProThrCysMetProIleAspCysGly 2614
Db 2170 GATGCTGAATAATGGAAATTTGGTTCAAAAAGATTCACCTTTGTAAAGTTATTCATCTGCAC 2229

```
QY 2615 LeuProHisIleAspPheGlyAspCysThrLysLeuLysAspGlnGlyTyrPhe 2634
DB 2230 CTTCCACCGAGTGTCAATGGGAG----- 2256
QY 2635 GluGlnGluAspMetMetGluValProTyrValThrProHisProTyrHisLeu 2654
DB 2257 -----CACACA 2262
QY 2655 GlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsn 2674
DB 2263 GGATGATGGCAGAA-----AAC 2280
QY 2675 PheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsn 2694
DB 2281 TTCTATATGGAATGAAGTCTCTATGATGTGACCAAGGATTCTATCTCTCTGGGAGAG 2340
QY 2695 ProValLeuLeuCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSer 2714
DB 2341 AAAAAATTGCAGTGCAGAGTGAATCTTAAAGGACATGGATCT----- 2382
QY 2715 IleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSer 2734
DB 2382 ----- 2382
QY 2735 MetGlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeu 2754
DB 2382 ----- 2382
QY 2755 ArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCys----- 2770
DB 2383 -----TGAGCGGGCGCTTCCCCACAGTGTCTTACGATCTCTCT 2418
QY 2771 GluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLys----- 2787
DB 2419 CTTGTGACTCGCTGC-----CCTAATCCAGAGTCAAAATGGGTACAGCTCAATAAA 2472
QY 2788 ---GlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrVal 2806
DB 2473 ACACATTCTGCATATTTCCCAATAGCATAGTGTGTGACTGCAATCTCTGGCTTCATC 2532
QY 2807 LeuAsnGlyThrGluArgArgThrCysGluAspLysAsnTrpAspGluAspGluPro 2826
DB 2533 ATGAATGGTAGTGCCTGATTTAGTGTGATGATGATGATGATGATGATGATGATGATG 2592
QY 2827 IleCysIlePro-----ValAspCysSerSerProValSerAlaAsnGlyGln 2843
DB 2593 ACTTGATGAAAAGAGCTTCTATAGGTGTCCACCTCCGCTAAGACCCCTAAGCGAAC 2852
QY 2844 ValArgGlyAsp-----GluTyrThrPheGlnLysGluIleGluTyrThrCysAsn 2860
DB 2653 CATACTGGTGGAAACATAGCTCGATTTTCTCTGGAATGTCAATCTCTGTACAGCTGTGAC 2712
QY 2861 GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer 2880
DB 2713 CAAGGCTACCTGCTGGTGGAGAGGCACTCTCTTTTGACACATGAGGGAACCTGGAGC 2772
QY 2881 GlyAlaThrProAspCysValProValArgCysValAlaThrProGlnLeuAlaAsnGly 2900
DB 2773 CAACCTGCCCTCATTTGTAAGAGGTAAACTGTAGCTCACCAGCAGATATG-----GATGGA 2829
QY 2901 ValThrGluGlyLeuAsp-----TyrGlyPheMetLysGluValThrPheHis 2916
DB 2830 ATCCAGAAAGGGCTGGAACCAAGGAAAATGTATCAGTATGGAGCTGTGTGTAATCTGGAG 2889
QY 2917 CysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsn 2936
DB 2890 TGTGAAGATGGGTATATGCTGGAAGGCACTCCCCAGAGCCAGTCCCAATCGGATCACCAC 2949
QY 2937 TrpAspAlaGluIleProLeuCysLys-----ProValAsnCysGly 2950
DB 2950 TGGAACTCCCTCCCTGGGGTTTGACAGATCCCGTTCACTGTCTCTCTTTTGTGGT 3006
```

RESULT 11

```
US-09-230-652-1
; Sequence 1, Application US/09230652A
; Patent No. 6537775
; GENERAL INFORMATION:
; APPLICANT: Tournier-Lasserre, Elisabeth
; APPLICANT: Joutel, Anne
; APPLICANT: Bousser, Marie-Germaine
; APPLICANT: Bach, Jean-Francois
; TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATION
; FILE REFERENCE: 03715.0048-00000
; CURRENT APPLICATION NUMBER: US/09/230,652A
; EARLIER FILING DATE: 1999-05-17
; EARLIER APPLICATION NUMBER: FR 96 09733
; EARLIER FILING DATE: 1996-08-01
; EARLIER APPLICATION NUMBER: FR 97 04680
; EARLIER FILING DATE: 1997-04-16
; EARLIER APPLICATION NUMBER: PCT/FR97/01433
; EARLIER FILING DATE: 1997-07-31
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8091
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(7041)
; OTHER INFORMATION: human ADNC No. 6537775ch 3
US-09-230-652-1
```

Alignment Scores:

Pred. No.:	4,87e-72	Length:	8091
Score:	1004.00	Matches:	651
Percent Similarity:	31.12%	Conservative:	234
Best Local Similarity:	22.89%	Mismatches:	931
Query Match:	5.03%	Indels:	1034
DB:	4	Gaps:	157

US-09-977-053-4 (1-3571) x US-09-230-652-1 (1-8091)

QY	988	CysArgProGlySerValLeuArg---	GlyArgMetCysValAsn-----	1001
DB	562	TGCGGGGTGGTGGAGCCCTGCGCCATGCGGCATGCGGCACCTGCTCAACACACACCTGGCTCTTC	621	
QY	1002	-----CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer---	1017	
DB	622	CGCTGCCAGTGTCCAGCTGCTACACAGGGCCACTA-----	669	
QY	1018	-----CysArgIleGlySerTyrGlnAspGluGluGlyGln	1029	
DB	670	GCCTGCTGCTGCGCCCTCACCATGCGTAACGGGGGCACCTGCAGGCAGAGTGGCGAC	729	
QY	1030	Leu-----GluCysLysLeuCysPro-----SerGlyMetTyrThrGluTyrIleHis	1045	
DB	730	CTCACTTACAGTGTGCTCTCTCTCTGGGTTTGAGGGTTCAGAAATTTGTAAGTG-----	783	
QY	1046	SerArgAsnIleSerAspCysLysAlaGln-----	1058	
DB	784	-----AACGTGACGACTGTCCAGGACACCGATGTCTCAATGGGGGACATGCTGGAT	837	
QY	1059	GlyThrTyrSerTyrSer-----	1067	
DB	838	GGGTCAACACCTATAACTGCCAGTGCCTCTGAGTGGACAGGCCAGTTCTTGCACGGAG	897	
QY	1068	ThrCysGluSerCysProLeu-----	1079	
DB	898	GACCTGAGTGTGCTCAGCTGCAGCCCAACCGCTGCCACCAATGGGGGTACCTGCTTCAAC	957	
QY	1080	LysPheGlySerArgSerCysLeu-----	1091	
DB	958	ACGCTGGTGGCCACAGCTGCTGTGTGTGTCAATGGCTGGACAGGTGAGAGCTGAGTCTG	1017	

Qy	1092	Asn	-----ThrSerThrValLysArgGlyAlaVal-----AsnIle	1101
Db	1018	AAATTCATGATGACTGTGCACACAGCGGTGTGCTTCCATGGGGCCACCTGCCATGCCACGCGGTG	1077	---
Qy	1104	SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys	1123	---
Db	1078	GCCTTCTTTACTGTGTGCTGCCCATGGCAAG-----ACTGCCTTCCTG---TGT	1125	---
Qy	1124	His	-----ProCysProArgAspTyrTyr-----Gln	1132
Db	1126	CACCTGGATGACGCTGTGTGTCAGCAACCCCTGCCACGAGGTGCTATCTGTGCACACAAT	1185	---
Qy	1133	ProAsnAlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAla	1152	---
Db	1186	CCGCTGAACGGCGCGGCATTGTC---ACCTGTCT---CCCGCTTCACCGGTGGGGCA	1239	---
Qy	1153	GlySerArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGlu	1172	---
Db	1240	TGTGCACGAGTGTGGACGAGTGTCTTATCGCGCGCCACCCCTCGCGAGCACTTGGGGCAGG	1299	---
Qy	1173	SerValValProProAlaSerLeu-----GlyHisIleLysLysArg	1186	---
Db	1300	TGCGTGAACACCGACGGGCTCTCTCTGTGCCAGTGCGGTGGTGTACACTGCACCTCGC	1359	---
Qy	1187	HisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisAsnSerGly	1206	---
Db	1360	TGTCGAGACCGAT-----GTCAACGAGTGTCTGTGGGGCCCTGCCGAACACGAGCC	1410	---
Qy	1207	ThrCys---GlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGly	1225	---
Db	1411	ACGTGCCTCGACCGCATAGGCCAG---TTCACTGTATCTGTATGGCAGGCTTCACAGGA	1467	---
Qy	1226	LeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGlyVal	1245	---
Db	1468	ACCTATTGCGAGGTGGACATTACGACAGTGTCCAGAGTAGCCCTGTGTCAACGGTGGGTC	1527	---
Qy	1246	CysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGlnArg	1265	---
Db	1528	TGCAAGGACCGAGTCAATGGCTTCAGCTGCACCTGCCCTCGGGCTTCAGCGGCTCCACG	1587	---
Qy	1266	CysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLysGlyIleCysVal	1285	---
Db	1588	TGTCAGCTGGACGTGGACGAATGCCACGACGCGCTGCAGGAATGCGCCAAATGCGTG	1647	---
Qy	1286	AspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHisCysGlu	1305	---
Db	1648	GACCAGCCCGATGGCTACGAGTCCGCTGTGCCAGGGCTTTGAGGCCACGCTGTGTGAT	1707	---
Qy	1306	ThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCysGluAspGln	1325	---
Db	1708	CGCAACGTGGACGACTGCTCCCTGACCCATGCG---CACCATGGTCCGTGCGTGGATGGC	1764	---
Qy	1326	ValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGlyLysAsn	1345	---
Db	1765	ATCCCGACGCTTCTCATGTGCCTGTGCTCTTCGTGTACACGGGCACACGCTGCGAGGCCAG	1824	---
Qy	1346	ValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGlyAlaAsn	1365	---
Db	1825	GTGACGCAATCCGACGCCCTTCGCCCATGSGCGCAATGCCCTAGACCTGCTGGAC	1884	---
Qy	1366	SerPheArgCysLysCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAsn	1385	---
Db	1885	AAGTACCTCTGCGCTGCTCTGGGACCAACAGGTGTGAACCTGCGAAGTGAACTATGAC	1944	---
Qy	1386	GluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyr	1405	---
Db	1945	GACTGTGCCAGCAACCCCTGCG---ACCTTTGGAGTCTGCGCGTATGCGCATCAACCCGCTAC	2001	---
Qy	1406	SerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSerThrGly	1425	---
Db	2002	GACTGTGTCTCCCAACCTGGCTTCACAGGGCCCCCTTTGTAACTGGAG-----	2049	---
Qy	1426	PheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAsp-GlyMetLe	1445	---

[illegible]

2622 GGACATCAATGACTGT---GACCCCAACCCATGCTGAACGGTGGCTGCTGCCAAGACGG 2678
1798 yHisSerGlyGluLeuThrValGlyAlaAlaValThrPheSerCysGlnGluG1 1818
2679 C-----GTGGGC-----TCCTTTTCTCTCTCC----- 2700
1818 yTyrGlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLe 1838
2701 ---TGCTCTCTGCTTTCGCGGCCCCACGATGCGCCGGAT----- 2739
1838 uileProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysI1 1858
2740 -GTGGAGTGGCTGCTGAGCAACCCCTCGGC-----CGGGCACCTGTAC 2783
1858 eGluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrTh 1878
2784 CGACCACTGGCC-----TCCTTCACCTGCACCTGCGCGCGGTAC-- 2826
1878 rLeuAlaGlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProPr 1898
2827 -----GGAGGCTTCCACTGCCGACAGGACCTGCC 2855
1898 oValCysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLe 1918
2856 CGACTGCAGCCCGAGCTCTGCTTC-----AATGGCGGACCTGTGT 2897
1918 userGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuG1 1938
2898 GAGCGGC-----GTGAACCTGTTGAGCTGCTGTGCGCTGCCCGGTAC-- 2940
1938 nGlyProSerIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCy 1958
2941 -----ACAGAGCCCATGCTGCACATCAGCAGACCCC----- 2973
1958 sHisLeuValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAs 1978
2974 ---TGCTCTCGCGCCCTGCTACACGGGGGCTGTCAGCGCGGCCCA 3020
1978 nPheThrPheArgAsnThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLe 1998
3021 CCTTGGCTTCGC-----TGCACCTGCTCGAGAGCTTCACGGGCCCGCGAGTG 3068
1998 uAspThrIle-----GluCysLeuAlaAspGlyLysTrpSerAr 2011
3069 CCAGACGCTGGTGGATTGGTCAGCCCGCCAGCTTGTCAAACGGGGTCTGCTCGTCCA 3128
2011 gSerAspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSe 2031
3129 GACTGGGGCTATTGCTT-----TGTC 3152
2031 rProGluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTy 2051
3153 CCTTGGATGAGCGGAGCTCTGT---GACATCCGAAGCTTGCTCCCTGCGAGGAGCGCGC 3209
2051 rSer---LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProPr 2070
3210 AGCCAGATCGGGTGGCTGGAGCAGCTGTGTACGGCGGGTGGCAGTGTGTGATGA 3269
2070 oGluGlyGlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSe 2090
3270 AGACAGC-----TCCCACTACTGC----- 3288
2090 rTyrSerIleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSerPh 2110
3288 ----- 3288
2110 eLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGlu----- 2125
3289 -GTGTGCCAGAGGC-----CGTACTGTAGTCCACTGTGTGAGCAGGAGGTGGACCC 3338
2126 -----CysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleG1 2140
3339 CTGCTTGGCCAGCCCTGCGCAGCATGGGGG----- 3369

QY 2140 nCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySe 2160
Db 3370 -----ACCTGCGCTGCTATATATGGGGGC----- 3393
QY 2160 rAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysG1 2180
Db 3394 -----TACATGTGTGAG-----TGCTTCTCTGG 3416
QY 2180 yGluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHi 2200
Db 3417 CTACATATGGTATTAACCTGTGAGCAGCAGTGCAGTGTGCTCCACGACCTGCCAGCA 3476
QY 2200 sProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrG1 2220
Db 3477 CGGGGTTTCATGC----- 3489
QY 2220 yArgIlePheGluSerGluValArgTyrGlnCysAsn-----ProGlyTyrLysSerVa 2238
Db 3490 -----ATTGACCTGTGGCCCGCTATCTCTGCTCCCTGCTCCCGCAGAACGCTGGGGT 3542
QY 2238 lGlySerProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMe 2258
Db 3543 G-----CTCTCGCAGATTAAATCAGGAT----- 3564
QY 2258 tCysValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyG1 2278
Db 3565 -----GACTGCGGCCCGCCCGCTGACTCAGGGGCCCGCTGCTTACA 3611
QY 2278 uAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValG1 2298
Db 3612 CAAT-----GSCACCTGCGTG-----GACTGTGTGG 3638
QY 2298 yAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCy 2318
Db 3639 TGGTTTCGCTGCACCTGTCCC-----CCAGGATA 3668
QY 2318 sMetProAlaLysCysProGluProProLeuLeuGlnLeuValLeuLysGluLe 2338
Db 3669 CACTGGTTTGGCTGC-----GAGCGACATCAATGAGTGTGCG----- 3708
QY 2338 uThrThrGluValGlyValThrPheSerCysLysGluGlyHisVal-----Le 2355
Db 3709 -----TCAGGTGCTGCACCGCGCACACACCGGAGCTGCT 3746
QY 2355 uGlnGlyPro---SerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPhePr 2374
Db 3747 GCAGGACCCAGCGGAGGTTTCGTTGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 3800
QY 2374 oValCysLysIleValLeu-----CysThrProProLeuIleSerPheGlyValPr 2392
Db 3801 TCGTGTGACAGTGTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3860
QY 2392 oIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGlyPh 2412
Db 3861 CGGTCTTAGCCCGGCTCTGGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3920
QY 2412 ePhe-LeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProL 2432
Db 3921 CTGGGTCTCGCTTGCAGC-----GGGTGGCGCGCTCTCTG 3956
QY 2432 euProGluCys-ValProValGluCysProGlnProGluGluIleProAsnGlyIleIle 2451
Db 3957 CCGGGAGCTGAGTGTGCGGTGGCGGTGCTCCATGTCAGCAGACGCGCCCGCGG----- 4007
QY 2452 AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyr-ThrCysLysProGlyPheG1 2471
Db 4008 -----GCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4032
QY 2471 uLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysPr 2491
Db 4033 -----TTGTCGGAGACCTCTCTGTCGCGAGCTTTCGCGGGTGCCTGCC 4070

5575 -----GCACGACTGACCGTACTGGCGGAGAGTGTCTTGACCTGGCTGCCCGTTAT 5625
3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAsp 3159
5626 GCCGTGCTCATGAGCGAAGCGGCTGCTG-----GATGCTGGGGCAGAC 5670
3160 ThrPheThrCysGlnLysAspGlyArgTyrPheProGluArgLysSerCysSerProLys 3179
5671 ACCAATGCCAGGACCACTCAGGCGGCAC-TCCCTGTCACAC---AGCTGTTCAC---AGC 5723
3180 LysCysPro-----LeuProGluAsnIleThrHisIleLeuValHisGlyAspPhe 3197
5724 CGATGCCAGGGTGTCTTCCAGATTCTCATCCGAACCGCTCTACAGA----- 5771
3198 SerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsn 3217
5772 -----CTTGGATGCCCGCATGGCATGGCTCAACGGCACTGAT 5810
3218 IleSerValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSer 3237
5811 CCTGGC----- 5816
3238 ProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyr 3257
5817 -----GGCCGCGCTGGCAGTAGAGGCGATGCT-----GGAGAGCTCAT 5855
3258 -----ThrPheGluSerThrIleIleTyrGln 3266
5856 CGCCAGCATGCTGATGATGCTGATGCTGGATGAGCTTGGGAATCAGCTTTACACTGGGC 5915
3267 CysGluProGlyTyrGluLeuGluGlyAsnArgGluValCysGlnGluAsnArgGln 3286
5916 TGC-----GGCTGTGAACAGCTGGAGGACACTTTGGCCCTGCTCAA-----AAA 5960
3287 TrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeu 3306
5961 TGGAGCCA-----ATAAGGACATGCGATAGCAAGGAGGAGACCCCTATTCTCTGGCG 6016
3307 AsnGlyLysAlaAspIleGluAsnArg-----ThrThrGlyProAsnValVal 3322
6017 CCGCGAGGGGAGCTATGAGGCTGCCAAGCTGCTGTGGACCACTTTGCCAACCGTG--- 6073
3323 TyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer----- 3335
6074 -----AGATCACCGACCACTGGACAGGCTGCGCGGG 6106
3336 -----GluAlaHisCysThrGluAsn-----GlyThrTrpSerHisProVal--- 3349
6107 ACGTAGCCAGGAGAGACTGCACCGAGACATCGTGGCGCTTGTGATCAACCCAGTGGGC 6166
3350 ProLeuCysLysPro-AsnProCysPro-ValProPheValIleProGluAsnAlaLeuL 3369
6167 CCGCAGCCCCCGGTCGCCACCGCTGGGGCTCTGTCTGTCTTCCAGGGGCTTCC 6226
3369 euSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGlu----- 3386
6227 TCCCTGGCTCAA-----GGCGCACAGTGGGGTCCAGAGAGGAGGCGGCCCGC 6280
3387 -----GlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThr 3404
6281 GGAAGGCGGCGCTGGGCGCGCAGGG----- 6306
3404 rPThrGlnThrSerAlaLysCysGluLysIleSerCys--GlyProProAla---HisVa 3422
6307 --CCCCGGGGGGGGGCAAGAGCTGACGTGGCTGCGCGGCGGCGGCTGCTGCTACAGCT 6364
3422 lGluAsnAlaIleAlaGlyValHisIleTyrGlnTyrGlyAspMetIleThrTyrSerCy 3442
6365 CGGTACAGCTGTGCGCGCTGATC----- 6389
3442 sTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThr 3462
6390 -----GCTGGACTCCCGCGGCGCTTTCGTGGGCGGCGGCTGCTTC-----CCCTG 6433

3462 pThrSerProPro-----IleCysArgAlaValCysArgPheProCysGlnAsnG1 3479
6434 GTGCTTCCCTTCAGGGGCGCTATGAGCTGCGACCTGCCACTGCTCAGTGTCTCTGGCAC 6493
3479 yGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCy 3499
6494 AGCTTGGTGGCGCAGGCGCGGAGCTCTAGGCGCGCAGCCCGCTCG-----AGGATGTG 6547
3499 sGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrG1 3519
6548 TACTCAGCTGGGCTGCTGTAACCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6604
3519 nCyAspCysPro-----ProGlyTyrThrGlySerArg-----Cy 3531
6605 CCCACCTGCTCCCTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6663
3531 sHisThrAlaValCysGlnSerProCysLeuAsnGlyLysCysValArgProAsnAr 3551
6664 CTCAACCCAGGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6723
3551 gCysHisCysLeuSerSer-----TrpThrGlyHisAsnCysSerArgLysArg 3567
6724 GGACATGCGGAGGAGTACCGGTGCTGGGCGCACACAGCAGCCCGCCCAAGG 6775

RESULT 12

US-09-484-970B-65
; Sequence 65, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 65
; LENGTH: 8257
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6426186 285840.2
; NAME/KEY: unsure
; LOCATION: 7391, 8247
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-65

Alignment Scores:
Pred. No.: 1,31e-70 Length: 8257
Score: 987.00 Matches: 645
Percent Similarity: 30.47% Conservative: 239
Best Local Similarity: 22.23% Mismatches: 939
Query Match: 4.94% Indels: 1083
DB: 4 Gaps: 157

US-09-977-053-4 (1-3571) x US-09-484-970B-65 (1-8257)

398 CysArgProGlySerValLeuArg---GlyArgMetCysValAsn----- 1001
562 TGGCGGTGGTGGAGCCCTGCGGCATGTTGGCAGCTGCTCCTCACACACACTGGCTCTTC 621
1002 -----CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSer--- 1017
622 CGCTGCCAGTGTCCAGCTGCTACACAGGCGCACTA-----TGTGAGAACCCC 669
1018 -----CysArgIleGlySerTyrGlnAspGluGlyGln 1029
670 GCGGTGCCCTGTGCGCCCTACCATGCGTAACGGGGGACACTGTCAGGAGAGTGGCGAC 729

QY 2392 oileProSerSerAlaLeuHisPheGlySerThrVallySerCysValGlyGlyPh 2412
 DB |||||
 DB 3861 CGTCTCTAGCCGGGTCCTGGGGTGGCTGACCTTACCTGTCTACTGTGCCACGCGGT 3920
 QY 2412 ePhe-LeuArgGlyAenSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProL 2432
 DB |||||
 DB 3921 CTGGGGTCCGGCTTGGGAGC-----GGGTGGCGGCTCTCTG 3956
 QY 2432 euProGluCys-ValProValGluCysProGlnProGluGluLeuProAsnGlyIleIle 2451
 DB |||||
 DB 3957 CGGGAGCTGCAGTGCCTGGGCTGCCATGCCAGCAGACGCCCGCGG----- 4007
 QY 2452 AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyr-ThrCysLysProGlyPheGl 2471
 DB |||||
 DB 4008 -----GGCGCGCTGGCTTGCCTCCCGCCAGGG----- 4032
 QY 2471 uLeuValGlyAenThrThrThrLeuCysGlyGluAenGlyHisTrpLeuGlyGlyLysPr 2491
 DB |||||
 DB 4033 -----TGTGCGGACCTCTCTGCGCAGCTTCCCGGGTGGCC 4070
 QY 2491 o-----ThrCysLysAlaIleGluCysLeuLysProLysGluIleLe 2505
 DB |||||
 DB 4071 GCGGGGCGCCAGCACCGCAGCTGCGCGCGCCCTCTG----- 4110
 QY 2505 uAenGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr----- 2519
 DB |||||
 DB 4111 -----CTCACGGGGCTCTCTGCGCGCCCGCGCGCTCGC 4145
 QY 2520 -----ValThrTyrSerCysAsnArgGlyPhe-----ArgLeuGluGlyProSe 2534
 DB |||||
 DB 4146 GCCTTCTCTCCGCTGCGCGCAGGCTGACCGGGCGCGCTTGGAGGCGCGCGC 4205
 QY 2534 rAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaI 2554
 DB |||||
 DB 4206 CGCGGCACCGCAGGTCTCG-----GAGGAGCGCGGCTGCCCGCGCGC 4247
 QY 2554 eHisCysAspSerProGlnProIleGluAenGlyPheValGluGlyAlaAspTyrSerTy 2574
 DB |||||
 DB 4248 CCCTGCCAG----- 4257
 QY 2574 rGlyAlaIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGl 2594
 DB |||||
 DB 4258 -----GCCAAGCGGGGACCA 4274
 QY 2594 nThrCysGluGluSerGlyTrpSerSerIleProThrCysMetProIleAspCysGl 2614
 DB |||||
 DB 4275 GCGCTCGACCGCGAG-----TGCAACAGCCCGAGCTGCGG 4310
 QY 2614 yLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPh 2634
 DB |||||
 DB 4311 CTGG-----GACGGCGCGACTGCTCGCTGAGGTGGCGAC----- 4347
 QY 2634 eGluGlnGluAspAspMetGluValProTyrValThrProHisProTyrHisLe 2654
 DB |||||
 DB 4348 -----CCCTGGCGGCA 4358
 QY 2654 uGlyAlaValAlaLysThrTrp-----GluAsnThrLysGluSerProAlaThrHi 2671
 DB |||||
 DB 4359 ATGCGAGGCGCTGCAGTCTGGCGCTTCTTCAACACAGCGCTGCGACCCCGCTGCGAG 4418
 QY 2671 sSerSerAsnPheLeuTyrGlyThrMetValSerTyr-----ThrCy 2685
 DB |||||
 DB 4419 CTGCGCGCGCTGCTCTTACGACACTTCAGCTCCACGCGGTGGCGGCGGCGCACTTG 4478
 QY 2685 sAsnProGlyTyrGluLeuLeu-----GlyAsnProValLeuI 2698
 DB |||||
 DB 4479 CAACCGGTGTACGAGAGTACTGCCCGACCACTTTTGGCGAGCGCGCTGCGACCGAGG 4538
 QY 2698 eCysGln-----GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCy 2717
 DB |||||
 DB 4539 CTGCAACAGCGAGGAGTGGCGCTGGGATGG-----CTGGATTG 4577
 QY 2717 s-----AspLeuProThrAlaProGluAenGlyPheLeuArgPheThr----- 2731

DB 4578 TGCCAGCAGAGTCCCGGCCCTCTGCTGCCCGCGGGTGGTCTCAGAGTGTCTGTGCC 4637
 QY |||||
 DB 2732 -----GluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAl 2750
 DB |||||
 DB 4638 GCGGAGCAGAGTACTGCTGCTTCCAGCGCGCAGCTTTCTTCAGCGCTCAGCGCCATCTCC 4697
 QY |||||
 DB 2750 aGlySerAspLeuArgLeuCysLeuGlu-----As 2760
 DB |||||
 DB 4698 CACCTCG-----CTGCGCTTCCGCTGCGCGCAGCGCAGGCCATGGTCTTCCCTTACCA 4754
 QY |||||
 DB 2760 nArgLysTrpSerGlyValAspSerProArgCysGluAlaIleSerCysLysLysProAsnPr 2780
 DB |||||
 DB 4755 CGGGCTAGTCTGCTGCTCCGAACCCCGGGCCGT-----CGGAGCTGGGCC 4802
 QY |||||
 DB 2780 oValMetAenGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGl 2800
 DB |||||
 DB 4803 CGAGGTGATCGCTCGGTA-----GTAAATGCTGGA 4832
 QY |||||
 DB 2800 uCysAspProGlyTyrValLeuAenGlyThrGluArgArgThrCysGlnAspAspLysAs 2820
 DB |||||
 DB 4833 GATTGACACCGGCTCTGCTGCTGCGCTGAGATGATCATCTGCTTCCCGATGCCCA 4892
 QY |||||
 DB 2820 nTrpAspGluAspGluProIleCysIleProValAspCysSerSerPro-----ProVa 2838
 DB |||||
 DB 4893 GAGCGCGCTGACTA-CTTGGGAGCGT--GTCAGCGGTGGAGCGCTGCTGCTTCCGT 4948
 QY |||||
 DB 2838 lSerAlaAenGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrTh 2858
 DB |||||
 DB 4949 ACCACTGCGGACGCTGCGGG-GGGAGCGCTGGAGC----- 4984
 QY |||||
 DB 2858 rCysAenGluGlyPheLeuLeuGluGlyAlaArgSerArgValCys-LeuAlaAenGlyS 2878
 DB |||||
 DB 4985 -----CTCCGAACCCAGCGTCCCGCTGCTGCCACTGCTAGTGGCG 5025
 QY |||||
 DB 2878 eTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuA 2898
 DB |||||
 DB 5026 GG-----CGCTGT-----CTTG 5037
 QY |||||
 DB 2898 laAenGlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysH 2918
 DB |||||
 DB 5038 CTGCTGCT-----TGTCTATGGTGGCCCGCGCAGCGCGACAGCACCCTCTGCTTC 5052
 QY |||||
 DB 2918 iGluGlyTyrIleLeuHisGlyValaProLys----- 2928
 DB |||||
 DB 5053 GTCTCGG-----TGTCTATGGTGGCCCGCGCAGCGCGACAGCACCCTCTGCTTC 5106
 QY |||||
 DB 2929 -----LeuThrCysGlnSerAspGlyAsnTrp-----AspAlaGluIleProLeuC 2944
 DB |||||
 DB 5107 CCTGAGGGCTTCTCTACTGCACAGGACGTGGCTCTGCTCACAAGGGCGCGCGGAACCC 5166
 QY |||||
 DB 2944 yLysProValAsnCysGlyProProGluAspLeuAlaHisGly-----PheProA 2961
 DB |||||
 DB 5167 GTGGGCGCAGAGCGCTGCGGCATGAAGAA-----CATGGCCAAAGGGTGAGAGCTG 5217
 QY |||||
 DB 2961 sngLysPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysL 2981
 DB |||||
 DB 5218 ATGG-----GGAGGT-----GGCCACAGAC 5238
 QY |||||
 DB 2981 euHisGlyAenSerSerArgArg-----CysLeuSerAenGlySerTrpSerGlyS 2998
 DB |||||
 DB 5239 TGGATGCACAGAGTGCACAGGCGCTAAAGGTAGAGGAGCCAGGATGGGG 5298
 QY |||||
 DB 2998 eSerProSerCys---LeuProCysArgCysSerThrProVal-----IleG 3013
 DB |||||
 DB 5299 GCTGAGGAGGCTGTGGATTGCCCTCAGTGGACTCAACACCATCTGTTGTTGCTGTGACATC 5358
 QY |||||
 DB 3013 luTyrGlyThrValAenGlyThrAspPheAspCysGlyLysAlaIa-----A 3029
 DB |||||
 DB 5359 CGGCTGGCACCAGCCATGGCCTGACCTGACACCCAGCGCGCAGATGCTGATGGCATG 5418
 QY |||||
 DB 3029 rgIleGlnCysPhe-----LysGlyPheLysLeuLeuG 3040

5419 GATGTCAATGTGGTGGCCAGATGGCTTACCCCGCTAATGTGGCTTCTCTGTGGG 5478
3040 lYleuSerGluIleThrCysGluAlaaspGly----- 3050
5479 GGGGCTCTGGAGCAATGCCAACTGAAGAGGATGAGGCAGATGACATCAGCTAGCATC 5538
3051 -----GlnTrpSerSerGlyPheProHisCysGluH 3061
5539 ATCTCCGACTGATCTGCCAGGGGGCTCAGCTTGGGGCAGGACTGACCTACTG-GCGA 5597
3061 iethrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSert 3081
5598 GACTGCTTTGCACCTGGCTGCGCGCTTATGCC-----GTGCTGATGACCAAGCGGCTGCT 5654
3081 rPlys-----GluAsnValIleThrTyrsrCysArgSerGlyTyValIleGlnGlys 3099
5655 GGATGCTGGGGCAGACACCAATGCCAGGACCACTCAGGCGGCA----- 5698
3099 erSerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrr-ProValCysGlu 3118
5699 -----CTCCCTGCACACAG-----CTGTCAAGCCGATGCCCGGGTGTCTT 5741
3119 Pro-----LeuSerCys----- 3122
5742 CCAGATTCTCATCGAAACCGCTTACAGACTTGGATGCCCGCATGGCAGATGGCTCAAC 5801
3123 -----GlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHis 3138
5802 GGCACCTGATCTGGCGCGCGCT-----GGCAGTAGAGGGCATGGTGGAGAGCTCAT 5855
3139 -----ThrTyrgluSerGluValLysLeuArg 3147
5856 CGCCAGCCATGCTGATGTCATGCTGTGATGAGCTTGGGAATATCAGCTTACACTGGGC 5915
3148 CysLeuGluGlyTyThrMetAspThrAspThrPheThrCysGlnLysAspGly 3167
5916 TGC-----GGCTGTGAACAACGTGGA 5936
3168 -----ArgTrpPhePro-GluArgIleSerCysSerProLysLy 3180
5937 AGCCACTTTGGCCCTGCTCAAAAATGGCCAAATAGGACATGACGATAGCAAGAGGA 5996
3180 sCysProLeu-----ProGluAsnIleThrHisIleValHi 3193
5997 GACCCCTTATCTCTGGCGCGCGCGAGGCAGCTATGAGCTGCCAGCTGTGTG-- 6054
3193 sGlyAspAspPheSerValAsnArgGlnValSer-----Valse 3206
6055 -----GACCATTGTGCCAACCGTGAGATCACCGACCACTGGACAGGCTGCCGGGA 6107
3206 rCysAlaGluGlyTyThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyTh 3226
6108 CGTAGCCAGGAGAGACTGCACAGGACATCGTGCCTGTGATCAACCCAGTGGGCC 6167
3226 rTrpGluProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSe 3246
6168 CGCAGCCCCCCC-----GGTCCCCCAGCGGCTGGGGGCTCTGCTCTGT----- 6210
3246 rProGluHisGlyPheValValGlySerLysTyThrPheGluSerThrIleIleTyrgl 3266
6211 -CTCCAGGGGCTTCTCTCCCTGCTCAAAAGGGGCACAGTGGGGTCCAAAGAAGACGAG 6269
3266 nCysGluPro-----GlyTyrgluLeuGluGlyAsnArgGluArgValCysGlnGl 3283
6270 GAGGCCCGCGGAAGCGGGCTGGGGCGCGCAGGGGCGCGGGGGCGGGGCAAGAGCT 6329
3283 uAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrProLe 3303
6330 GACGCTGGGCTCGCGGGGCGGCTGTGCTGACAGCTCGGTCAAGCTG-----TCGCCCGT 6383
3303 uGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnValValTy 3323
6384 GGACTGCTGGACTCCCCCGGGCTTTTGGTGGGGCGGCTTCTCCCTGTGT----- 6435

3323 rSerCysAsnArgGlyTyrsrLeuGluGlyPro-----SerGluAlaHisCysTh 3340
6436 -----GGCTTCCCTTGGGGGCGCTTATGACAGTCCAC-TGCCACTGCAG 6481
3340 rGluAsnGlyThr-----TrpSerHisProValProLeuCysLysProAsnPro----- 3356
6482 TGTCTCTGGCACAGCTTGGTGGCCAGCGCGGGGTCTAGGGCGCGCCCTCTGGGA 6541
3356 ----- 3356
6542 GATGTGTACTCAGCCCTGGGCTCTGTAACCTGTGTGCTGTGCCCCCTGATGTGGCCCGG 6601
3357 -CysPro-----ValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTy 3375
6602 CTGCCCCCAGCTGCCCTCCAGGCGCTCTCTCTCTGTC----- 6638
3375 rValAspGlnAsnValserIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyI 3395
6638 ----- 6638
3395 elleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSe 3415
6639 -----TGGCAC-----TGGCGCGGGACCCC-----AG 6661
3415 rCys-----GlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTy 3432
6662 CTGCTCAACCCAGGACCCCGCTCCCCGAG----- 6695
3432 rGlnTyrglyAspMetIleThrTyrsrCysTyrsrGlyTyMetLeuGluGlyPheLe 3452
6696 -----AGCGGCGCGCGCTT----- 6710
3452 uArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCy 3472
6711 -----ACCTGGCAGTCCCGACATGGCGAGGATACCCG 6745
3472 sArgPheProCysGluAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGl 3492
6746 GCGGCTG-----GGGCACACAGCAGCCCCCAAGCGCGCTTCTTGGGGTCCAGT 6799
3492 uGlyTrpMet-----GlyArgLeuCysGluGluProIleCysIleLe 3506
6800 GAGCACCTTACCTGACCCCATCCCGAATCCCTGAGCAGCTGGGCGCAGCCCTCACCT 6859
3506 uProCysLeuAsnGlyGly-----ArgCysValAlaProTyrglnCysAspCysPr 3523
6860 CCTCCTCTCTCAGACTGGTCCGAATCCACGCTAGCCCGCA-----CTGCC 6907
3523 oProGly-----TrpThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuAs 3542
6908 ACTGGGGCCATGGCCACCACTGGGGGCACTGCTGCCAGCCACTTCTCTGTCTGT 6967
3542 nGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrp 3558
6968 CCCAGCTCCCTGCTCAGGCCAG-----CCAGCTGG 7001

RESULT 13

US-08-083-590A-21

Sequence 21, Application US/08083590A

Patent No. 5786158

GENERAL INFORMATION:

APPLICANT: Artavanis-Tsakonas, S. et al.

TITLE OF INVENTION: Therapeutic And Diagnostic Methods

TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And

TITLE OF INVENTION: Nucleic Acids

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSER: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.


```

; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
; FILING DATE: 25-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..7419
; US-08-083-590A-21

Alignment Scores:
Pred. No.: 9,768-70 Length: 9723
Score: 978.00 Matches: 636
Percent Similarity: 30.26% Conservative: 263
Best Local Similarity: 21.41% Mismatches: 881
Query Match: 4.90% Indels: 1196
DB: 1 Gaps: 167

US-09-977-053-4 (1-3571) x US-08-083-590A-21 (1-9723)
QY 990 ProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTyrAsn 1009
DB 571 CCAGGACACTGCCAGATGGTGGACCTGCTCCCTCAACCTGCCT- 612
QY 1010 LeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGluGln 1029
DB 613 -----GGTTCTCTACAG 624
QY 1030 LeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAsnIle 1049
DB 625 -----TGCCAG-----TGCCCTCAGGCTTCACAGCCAGTACTGTGCACAGCTGTATGTG 675
QY 1050 SerAspCysLysAlaGlnCys---LysGlnGlyThrTyrSerTyrSerGlyLeuGluThr 1068
DB 676 CCTGTGCACCTCCCTTGTGTCATGAGGAGCACCCTGTGCGCAGACTGCTGACTTCACT 735
QY 1069 CysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySerArgSerCysLeuSer 1088
DB 736 TTTGAGTGCAACTGCTTCCAGGTTTGTGAAGGAGCACCCTGTGAGAGGAATTTGATGAC 795
QY 1089 CysProGluAsnThrSerThrValLysArgGlyAlaVal-----AsnIleSerAla 1105
DB 796 TGCCCTTACCAC-----AGGTGTCAAGATGAGGGGTTTGTGTGATGGGCTCAACACT 849
QY 1106 CysGlyValProCysPro-----GluGlyLysPheSerArgSerGlyLeuMetPro 1122
DB 850 TACAACTGCGGCTGTCCCCACAAATGGACAGACAGTCTGTGCACAGAGATGTGGATGAA 909
QY 1123 CysHisProCysProArgAspTyrTyrGlnProAsnAla----- 1135
DB 910 TGC-----CTGTGCGAGCCCAATGCTGTCAAAATGGGGGCGCCTGT 951
QY 1136 -----GlyLysAlaPheCysLeuAla----- 1142

```

```

1011 GCGAAGCGCAATGGAGGCTATGCTGTATGTGTCAACGGCTGGAGTGGAGTACTGTC 1011
1143 -----CysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
1012 AGTGAGAACATTGATGATTTGCTTCCCTTCTGCTACTTCA-----GGCTCCACCTGC 1065
1157 IleThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSerValPro 1176
1066 ATGACCGGTGTGGCTCTCTCTCT-----TCATGTGCGCCA 1101
1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
1102 GAGGGGAAGGCAGGT-----CTCCTGTGTCTATCGGATGATGCA 1140
1197 CysPhePheAsnProCysHisAsnSerGlyThrCys-----GlnGlnLeuGlyArgGly 1214
1141 TGCATCAGCAATCTTGGCCACCAAGGGGCGACTGTGTGACACCAACCCCTTAATGGGCNA 1200
1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGlu 1234
1201 TATATTTGCACCTGCCACAGCTTACAAGGGGCTGACTGCACAGAGATGTGGATGAA 1260
1235 CysSerProLeu-----ProCysLeuAsnAsnGlyValCysLysAspLeuValGly 1251
1261 TGTGCCATGGCCCAATAGCAATCTTGTGTGAGCATGACAGGAAATGTGTGAACACGGATGC 1320
1252 GluPheIleCysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsn 1271
1321 GCCTTCCACTGTGAGTGTCTGAAGGGTTATGACAGGACCTGCTGTGTGAGATGGACATCAAT 1380
1272 GluCysSerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyr 1291
1381 GAGTGCCATTGACCCCTGCCAGATGATGATCTACTCTGTCTGGATTAAGATGGAGGCTTC 1440
1292 ArgCysThrCysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCys 1311
1441 ACATGTCTGTGCATGCCAGGTTTCAAGGGTGTGTCATTGTGAATTAGAAATAAATGAATGT 1500
1312 GlnSerAsnProCysLeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCys 1331
1501 CAGAGCAACCTTGTGTGAACAAATGGGCAGTGTGTGGATAAAGTCAATCGTTTCCAGTGC 1560
1332 LysCysProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSer 1351
1561 CTGTGTCTCTCTGTTTCACTGCGCCAGTTTGCAGATTTGCCAGATTGATATTGATCTTCCAGT 1620
1352 GlnProCysLysAsnGlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCys 1371
1621 ACTCCGTGTCTGAATGGGCAAGTGTATCGATCACCCGAAATGGCTATGAATGCCAGTGT 1680
1372 AlaAlaGlyPheThrGlySerHisCysGlyLeuAsn----- 1383
1681 GCCACAGGTTTCACTGCTGTGTGTGTGTGAGGAGAACATTTGACAACTGTGTGACCCCGATCCT 1740
1383 ----- 1383
1741 TGCCACCATGGTCACTGTGCGAGTGTATTGATTCTCTACACCTGCATCTGCAATCCCGGG 1800
1384 -----IleAsnGluCysGlnSerAsnProCysArgAsn 1394
1801 TACATGGGGGCCCATCTGCAGTGCACAGATTGATGAATGTTTACAGCAGCCCTTGCCTGAAC 1860
1395 GlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSer 1414
1861 GATGTGCGCTGATTTGACCTGCTCAATGGCTACCTGCTGAGTGAACCTGTCAGCCGACGCTCA 1920
1415 GlyGlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
1921 GGGGTAAATTGTGAA-----ATTAATTTTGTGATGTGTGCAAGTAAC 1962
1435 -----IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThr 1452

```


QY 2053 uAlaasp-----AsnSerGlnLeuLeuCY 2061
DB 3612 TATTGACCTTGTGAACCACTTCAAGTGTCTTGGCCACAGGCACTCGGGGCTACTCTG 3671
QY 2061 sAsn-----AlaGlnGlyLysTrpValProGluGlyGlnAspMe 2075
DB 3672 TGAAGAGAACATTGATGACTGTGCCGGGGTCCCATTTAGCTTAAATGTTGTGTCAGTGCAT 3731
QY 2075 tProArgCysIleAlaHisPheCysGluLysProProSerValSerTrpSerIleLeuG1 2095
DB 3732 GCATAGCATTTGGAGGCTACAGTTGT----- 3756
QY 2095 uSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluG1 2115
DB 3757 -----CCTGCTTTCCTCTGG 3770
QY 2115 yPheVal-----LeuAsnThrSerAlaLysIleGluCysMetAr 2128
DB 3771 CTTTGTGGGAGCGTTGTGAGGGAGACATCAAC-----GAGTGCCTCTC 3815
QY 2128 gGlyGlyGlnTrpAsnPro-----SerProMetSerIleGlnCysIleProValArgCy 2146
DB 3816 C-----AACCCCTGCGAGCTCTGAGGGCAGCCTGAGCTGTATA----- 3852
QY 2146 sGlyGluProProSerIleMetAsnGlyTyAlaSerGlySerAsnTyrSerPheGlyAl 2166
DB 3853 -----CAGCTCACCAATGACTACTGTGT----- 3876
QY 2166 aMetValAlaTyrSerCysAsnLysGlyPheTyIleLysGlyLysLysSerThrCy 2186
DB 3877 -----GTTTGGCGTAGTGCTTT-----ACTGGCCGCACTG 3908
QY 2186 sGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisProValSerCysGlyG1 2206
DB 3909 TGAACCTTTCGTCGATGTGTGCCAGATGCCCTGCCATGAGGAGACTTGTGCTGT 3968
QY 2206 uProProLysValGluAsnGlyPheLeuGluHisThrGlyAlaGlyPheGluSerG1 2226
DB 3969 GGCCAGTACATGCTCTGATGGTTTCATT----- 3996
QY 2226 uValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPheValCysG1 2246
DB 3997 -----TGCCGTTGTCCCGGGATTTTCGGGGGCAAGG-----TGCCA 4034
QY 2246 nAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValProLeuAspCysGlyLys 2266
DB 4035 GAGCAGC-----TGTCGA-- 4047
QY 2266 sProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysVa 2286
DB 4047 ----- 4047
QY 2286 lGlnPhePheCysAsnGluGlyTyxGluLeuValGlyAspSerSerTrpThrCysGlnLys 2306
DB 4048 -CAAGTGAATGTAGGAAGGGGAGCAGTGTG----- 4080
QY 2306 sSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluPr 2326
DB 4081 -----CACACCGCTCTGACCCCGCTGCTTC-----TGCCCGAGTCC 4118
QY 2326 oProLeuLeuGluAsnGlnLeuValLysGluLeuThrThrGluValGlyValValTh 2346
DB 4119 C-----CGGGACTGCGAGTCAAGCTGTCAG 4145
QY 2346 rPheSerCysLysGluGly-----HisValLeuGlnGlyProSerValLeuLysCy 2363
DB 4146 TAGCCCTGCCAGCAGCGGGGAGCTGCGCACCTCAGCGCCAGCTCTTATTACTCCTG 4205
QY 2363 sLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCysThrPr 2383
DB 4206 C-----CAGTGTGCCCCACCATTCGTGGGTAGCCGCTGTGAACCTCTACACGGC 4253
QY 2383 oProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGlySerTh 2403

DB 4254 ACCCCCC-----AGCAC 4265
QY 2403 rValLysTyrSerCysValGly-----GlyPhePheLe 2414
DB 4266 CCTCTCTCCACCTGTCTGAGCCAGTATTGTGCCGCAAAAGCTCGGGATGGGCTCTGTGA 4325
QY 2414 uArgGlyAsnSerThrThrLeuCysGlnProAspGly-----Thr-Tips 2429
DB 4326 TGAGGCTGTCACAGCCATCGCTCCAGTGGATGGGGTGACTGTTCTCTCACCATTGA 4385
QY 2429 erSerProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsnG 2449
DB 4386 GNAACCTTGGGCCCACTGCT-----CCTCCCCACTTCCCTGCTG 4424
QY 2449 lYlleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProG 2469
DB 4425 GGAATTATCAACCAACCAAGTGTGTGATGAGCTGTGCAACACGGTCGAGTGCCTGTTT-GACA 4483
QY 2469 lYpHeGluLeuValGlyAsnThrThr-----LeuCysGlyGluAsnG 2484
DB 4484 ACTTTGAATGCCAGGGAGCAAGACATGCAAGTATGACAAATACTGTGCGACACCAC- 4542
QY 2484 lYHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLysProLysGluI 2504
DB 4542 ----- 4542
QY 2504 lLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerC 2524
DB 4543 -----TTCAAAGACAACAC- 4558
QY 2524 yAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspT 2544
DB 4559 GTAACCAAGGTGCAACAGTGA-----GAGTGTGCT--T 4591
QY 2544 rPAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluA 2564
DB 4592 GGGAT-----GGCTGGACTGTCTGTCTGACCAACCT--GAGA 4627
QY 2564 sNGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPheP 2584
DB 4628 AC-----CTGGCAAGAGGTACCTGTTATTGTGGTATTGATGTC 4666
QY 2584 rGlyPheGlnValAlaGlyHisAlaMetGln-ThrCysGluGluSerGlyTrpSerSer 2603
DB 4667 CACCTGACAACTGTCTCCAGGATGTCTGCGACTTCTTC-----GGGCATCG----- 4714
QY 2604 SerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPheGlyAsp 2623
DB 4715 -----GTACCCTGCTCCACACCA-----ACC 4735
QY 2624 CysThrLysLeuLysAspGlnGlyTyrPheGluGlnGluAspAspMetMetGluVal 2643
DB 4736 TCGGCATT-AAGCGGAGCTCCAGGG-----GAACCTCATGTGTGTAC 4776
QY 2644 ProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsn 2663
DB 4777 CCTATTATGCT-----GAGAAGTCAGCTGCTATGAAGAAACAGAGATG 4821
QY 2664 ThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyr 2683
DB 4822 ACACGCAGATCC----- 4833
QY 2684 ThrCysAsnProGly-----TyrGluLeuLeuGlyAsnProValLeuIle----- 2698
DB 4834 -----CTTCTCTGGTGAACAGACAGGAGGTGGTGGCTCTTAAGTCTTCTGGAAATT 4887
QY 2699 -----CysGlnGluAspGlyThrTrp-----AsnGlySerAlaProSer 2711
DB 4888 GACAACCGCCAGTGTGTTCAAGACTCAGACCACTGCTTCAAGAACACGGATGCGACGCA 4947
QY 2712 CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
DB -----

QY 3360 opheValIleProGlu---AsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAs 3379
DB 6605 CTTTCTTAACTTCATGAATGCACGCTTGGC----- 6638
QY 3379 nValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAs 3399
DB 6638 ----- 6638
QY 3399 nProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPr 3419
DB 6639 -----ACATGGGGCCAGCACTGTGCTTCCCTCAGTGAGCCAGTGTGCTATCCACC 6688
QY 3419 oAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleTh 3439
DB 6689 ACCACAT-----TGTGTCTCCAG-----CA 6709
QY 3439 rTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGlu-- 3458
DB 6710 GTGGCAGTCTGGAG-----CTTGAGTA 6733
QY 3458 ----- 3458
DB 6734 GGCTCCATCCAGTCCAGTCCAGTCCAGATGGATGAACCGCATGGAGTGAATGAGACCC 6793
QY 3459 -----AsnGlyThrTrpThrSe 3464
DB 6794 AGTACAAAGAGATGTTTGGTATGGTCTCGCTCCAGTGAAGGCACCCATCTGGCATAG 6853
QY 3464 rProPro-----IleCysArgAl 3470
DB 6854 CTCGCCAGAGAGCCACCTGAAGGAGACACATACACCCCTCGGAGCCCTGCCCC 6913
QY 3470 aValCysArgPhePro-----CysGlnAsnGlyGly----- 3480
DB 6914 CCATTGTGACTTCCAGCTCATCCCTAAAGGCGAGTATGCCCCAACAGCGGGGCTCCCC 6973
QY 3481 ----IleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCy 3499
DB 6974 AGCTCAGTCCACCTCCCTCCAGCTGT-----TG 7003
QY 3499 sGluGluProIleCysIleLeuProCysLeuAsnGlyGly-----ArgCysValAl 3516
DB 7004 CGGGCCCCCTGCCACCATGATCCAGATCCAGAAATGGCCGTTTGGCCAGTGTGGCTT 7063
QY 3516 aProTyrGlnCysAspCysProProGlyTrpThrGlySerArgCysHisThrAlaValCy 3536
DB 7064 TCCCACTGCCATGATGCCCGAGCAGGCGGCGAGGTAGTCTCAGACCATTCCTCCAGCT 7123
QY 3536 sGlnSer---ProCysLeuAsnGly 3543
DB 7124 ATCATCTTTCCAGCCCTCTGTGGG 7148

RESULT 14

US-08-532-384-21
; Sequence 21, Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,384
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/083,590
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCES/DOCKET NUMBER: 7326-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 9723 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 10..7419
US-08-532-384-21

Alignment Scores:
Pred. No.: 9,76e-70 Length: 9723
Score: 978.00 Matches: 636
Percent Similarity: 30.26% Conservative: 263
Best Local Similarity: 21.41% Mismatches: 881
Query Match: 4.90% Indels: 1196
DB: 3 Gaps: 167

US-09-977-053-4 (1-3571) x US-08-532-384-21 (1-9723)
QY 990 ProGlySerValLeuArgGlyMetCysValAsnCysProLeuGlyThrTyrTyrAsn 1009
DB 571 CCAGGACACTGCCAGCATGTGGCAGCTGCCTGCCT----- 612
QY 1010 LeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGluGln 1029
DB 613 -----GGTCTCTACCAG----- 624
QY 1030 LeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAsnIle 1049
DB 625 -----TGCAG--TGCCCTCAGGCGCTTCACAGCGCAGTACTGTGACAGCCTGTATGTG 675
QY 1050 SerAspCysLysAlaGlnCys---LysGlnGlyThrTyrSerTyrSerGlyLeuGluThr 1068
DB 676 CCCTGTGCACCTCCCTGTGTGTCAATGAGGAGCAGCTGTGCGCAGAGCTGTGATCACT 735
QY 1069 CysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySerArgSerCysLeuSer 1088
DB 736 TTTGAGTGCACCTGCTTCCAGGTTTGAAGGAGGAGCAGCTGTGAGAGGATATTGATGAC 795
QY 1089 CysProGluAsnThrSerThrValLysArgGlyAlaVal-----AsnIleSerAla 1105
DB 796 TGCCTTAACCCAC-----AGGTGTCAAGATGGAGGGGTTTGTGTGATGGGTCACACT 849
QY 1106 CysGlyValProCysPro-----GluGlyLysPheSerArgSerGlyLeuMetPro 1122
DB 850 TACAACCTGGCGCTGTCCCCACATGAGCAGGACAGTTCGACAGAGGATGTGGATGAA 909
QY 1123 CysHisProCysProArgAspTyrTyrGlnProAsnAla----- 1135
DB 910 TGC-----CTGCTGCAGCCCAATGCTCTCAAAATGGGGGACCTGT 951
QY 1136 -----GlyLysAlaPheCysLeuAla----- 1142
DB 952 GCCAACCCGCAATGGAGGCTATGGCTGTGTATGTGTCAACCGGCTGGAGTGGAGATGACTGC 1011

Db	2728	TGTGAGGAGGACATTGATGATCTGCCTTGGCAATCTTCCAGAAATGAGAGTTCCTGATAG	2787	3612	TATTGACCTTGTGACCACTTTCAAGTGTCTTTGCCACCAGGACACTCGGGGCCCTACTCTG	3671
Qy	1785	GLuProfile	1790	2061	sAbn	2075
Db	2788	GATGAGGAGTAATCTTCTCTGCTGCTCTGCTTCCGGGTTTCACTGGGAGTAAGTGCAG	2847	3672	TCAAGAGAACATTTGATGACTGTGCCCGGGTCCCATTCCTTAATGGTGGTCAGTGCAT	3731
Qy	1791	AlaProGlyAsnPro	1804	2075	tProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuG1	2095
Db	2848	ACAGACATGAATGAGTGTCTGAGTGAACCTGTAAGATGGAGGACCTGCTCTCACTAC	2907	3732	GGTAGGATTGGAGCTACAGTTGT	3756
Qy	1805	TyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyVal	1824	2095	uSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluG1	2115
Db	2908	GTCAAC	2952	3757		3770
Qy	1825		1833	2115	yPheVal	2128
Db	2953	CATTGTGAGAACACATCAATGAGTGCACTGAGAGCTCTCTTCAATGGTGGCACATGT	3012	3771	CTTTGCTGGGAGCCTTGTGAGGGAGACATCAAC	3815
Qy	1834		1842	2128	gGlyGlyGlnTyrAsnPro	2146
Db	3013	GTTGATGGGATTAATCTCTTCTGCTTGTGCTTGTGGCTTTCACCTGGATCTCTTGC	3072	3816	C	3852
Qy	1843		1859	2146	sGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAl	2166
Db	3073	CTCCATGAGATCAATGAATGCAGCTCTCATGCCCTGAATGAGGAGACGTGTGTGAT	3132	3853		3876
Qy	1860	GluLeuAlaPheThrPheGlySerLysValThrTyrArgCys	1873	2166	aMetValAlaLysSerCysAsnLysGlyPheTyrIleLysGlyLysLysSerThrCy	2186
Db	3133	GGCCTGGGT	3171	3877		3908
Qy	1874		1875	2186	sGluAlaThrGlyGlnTyrPheSerProIleProThrCysHisProValSerCysGlyG1	2206
Db	3172	ACTGGGAAAACTGTCAGACCCCTGGTGAATCTTCGAGTCGGTCTCCATGTAAAAACAA	3231	3909	TGAAACCTTCGTGATGTGTGCCAGATGCTCCCTGTAATGGAGGAGCTTGTGCTGT	3968
Qy	1876	GlyTyrThrLeuAlaGlyAspLysGluSerCysLeuAlaAsnSerSerTyrSerHis	1895	2206	uProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerG1	2226
Db	3232	GGTACTTGTGTCAGAAAAACAGACAGTCCAGTGCCTATGTCCATCTGGATGGCTGT	3291	3969	GGCCAGTAACATGCTGATGTTTCAAT	3996
Qy	1896	SerProProValCysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLys	1915	2226	uValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPheValCysG1	2246
Db	3292	GCC	3294	3997		4034
Qy	1916	TyrIleLeuSerGlyLeuThrTyrIleuSerThrAlaSerTyrSerCysAspThrGlyTyr	1935	2246	nAlaAsnArgHisTyrHisSerGluSerProLeuMetCysValProLeuAspCysGlyLy	2266
Db	3295		3333	4035	GAGCAGC	4047
Qy	1936	SerLeuGlnGlyProSerIleGlu	1953	2266	sProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysVa	2286
Db	3334	TCCAGGAGAGGT	3390	4047		4047
Qy	1954	AlaProProAlaCysHisLeuValPheCys-GlyGluProProAlaIleLysAspAlaVa	1973	2286	IcInPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLy	2306
Db	3391	GCTGCC	3441	4048	-CAAGTGAATATAGAGAGGGGAGCAGTGTGTG	4080
Qy	1973	IleThrGlyAsnAsnPheThrPheArgAsnThrValThrTyrThrCysLysGluGlyTy	1993	2306	sSerGlyLysTyrAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluPr	2326
Db	3442	TGTGAGGAGCACTCGATGAGTGTGTCCTCAACCCCTGCCAGACGGGCAACATGCAGT	3501	4081		4118
Qy	1993	rThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTyrSerArgSerAs	2013	2326	oProLeuLeuGluAsnGlnLeuValIleuLysGluLeuThrThrGluValGlyValTh	2346
Db	3502	GACTTC-ATTGTGGATACAGATGAGTGTGTCCAGGCTATCAG	3546	4119	C	4145
Qy	2013	pGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerProG1	2033	2346	rPheSerCysLysGluGly	2363
Db	3547		3587	4146	TAGCCCCCTCCACGACGCGGGGAGCTGCCACCCCTCAGCCGACCTCTTATTACTCTG	4205
Qy	2033	uThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLe	2053	2363	sLeuProSerGlnGlnTyrAsnAspSerPheProValCysLysIleValLeuCysThrPr	2383
Db	3588	GCCC	3611	4206	C	4253
Qy	2053	uAlaAsp	2061	2383	oProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGlySerTh	2403
				4254	ACCCCCC	4265

QY 2403 rVallyTySerCysValGly-----GlyPhePheLe 2414
 DB CCTCTGCGCACCTGTGTGAGCCAGTATTGTGCGCAAAAGCTCGGAGGGCTGTGTGA 4325
 QY 2414 uArgGlyAsnSerThrLeuCysGlnProAspGly-----Thr-Tips 2429
 DB TGAGGCTGCAACAGCAGCATGCTGCGAGTGGGATGGGGTGACTGTCTCTCACATGGA 4385
 QY 2429 erSerProLeuProGluCysValProValGluCysProGlnProGluGluLeuProAsnG 2449
 DB GAACCCCTGGGCCAACTGCT-----CCTCCCACTTCCCTGCTG 4424
 QY 2449 lYleileaspValGlnGlyLeuAlaTyLeuSerThrAlaLeuTyThrCysGlyProG 2469
 DB GGATTTATATCAACCAACAGGTGTGATGAGTGTGCAACACGGTGCAGTGTGTTT-GACA 4483
 QY 2469 lYpHeGluLeuValGlyAsnThr-Thr-----LeuCysGlyGluAsnG 2484
 DB ACTTTGATGCCAGGGGACAGCAGACATGCAAGTATGCAATATACTGTGCAGACCAC- 4542
 QY 2484 lYHisTrpLeuGlyGlyLysProThrCysLysAlaileGluCysLeuLysProLysGluI 2504
 DB ----- 4542
 QY 2504 leLeuAsnGlyLysPheSerTyThrAspLeuHisTyGlyGlnThrValThrTy-SerC 2524
 DB -----TTCAAGACAAACAC-----T 4558
 QY 2524 yAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspT 2544
 DB GTAACCGGGTCCACAGTGA-----GAGTGTGGT-----T 4591
 QY 2544 rPaspValAspAlaProSerCysAsnAlaileHisCysAspSerProGlnProileGluA 2564
 DB CGGAT-----GGGCTGACTGTGCTGTGACCAACCT---GAGA 4627
 QY 2564 snGlyPheValGluGlyAlaAspTySerTyGlyAlailellelleTySerCysPheP 2584
 DB -----CTGGCAGAGGTACCTGGTTATTGTGTATGATGTC 4666
 QY 2584 roGlyPheGlnValAlaGlyHisAlaMetGln-ThrCysGluGluSerGlyTrpSerSer 2603
 DB CACCTGAACACATGCTCCAGGATGCTCGCAGCTTCTTGC-----GGGCACTGG----- 4714
 QY 2604 SerileProThrCysMetProileAspCysGlyLeuProProHisileaspPheGlyAsp 2623
 DB -----GTACCCTGCTCCACACCA-----ACC 4735
 QY 2624 CysThrLysLeuLysAspAspGlnGlyTyPheGluGlnGluAspMetMetGluVal 2643
 DB TGGCGAAT-AAGCGGACTCCCGAGGG-----GAATCATGTTGTGATC 4776
 QY 2644 ProTyValThrProHisProTyHisLeuGlyValAlaValAlaLysThrTrpGluAsn 2663
 DB CCTATTATGCT-----GAGAAGTCAGCTCTATGAAGAACAACAGAGGATG 4821
 QY 2664 ThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyThrMetValSerTy 2683
 DB ACACGAGATCC----- 4833
 QY 2684 ThrCysAsnProGly-----TyGluLeuLeuGlyAsnProValleulle----- 2698
 DB -----CTTCTCTGGTGAAACAAGAACAGGAGGTGGCTCTTAAGTCTTCTTGGAAAT 4887
 QY 2699 -----CysGlnGluAspGlyThrTrp-----AsnGlySerAlaProSer 2711
 DB GACAACCGCCAGTGTCTCAAGACTCAGACCCTGCTTCAAGAACACAGGATGCGAGCA 4947
 QY 2712 CysileSerileCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
 DB -----CTGGCTCTCAGCGGCATACAGGGGACCTGTCTATAC--- 4989

QY 2732 GluThrSerMetGlySerAlaValGlnTy-SerCysLysProGly----- 2746
 DB -----CCTCTGTGTCTGTGTCAGTGAATCCCTGACTCCAGAACGCACTCAGCTCTC 5043
 QY 2747 HisileLeuAlaGlySerAsp-----LeuArgLeuCysLeu 2758
 DB TATCTCTCTCTGT 5103
 QY 2759 GluAsnArgLys-----Tip-----SerGlyAlaSerProArgCys 2770
 DB GCAAAACGAAAGCGTAAGCATGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5163
 QY 2771 GluAlaileSerCysLysLysProAsnProValMetAsnGlySerileLeLysGlySerAsn 2790
 DB ----- 5223
 QY 2791 TyThrTyThrLeuSerThrLeuTyTyTyGlyCysAspProGlyTyTyValLeuAsnGlyThr 2810
 DB -----CTCTCAGTCAAGTCTCAGAACCTAAC-----CTAATTGGTACT 5262
 QY 2811 GluArgArgThrCysGlnAspAspLysAsnTrpAspGluAspGluProileCysilePro 2830
 DB -----GGAACACAGTGAACACTGGGTGATGATGAA----- 5292
 QY 2831 ValAspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTyThr 2850
 DB -----GGGCCCCAGCCAAAGAAAGTAAGGCTGAA-----GATGAGCCTTA 5334
 QY 2851 PheGlnLysGluLeuGlu-----TyThrCysAsnGluGlyPheLeu 2864
 DB CTCTCAGAACAGATGACCCCATTTGATCGACGGCCATGGACACAGCAGCAC- 5385
 QY 2865 LeuGluGlyAla-----ArgSerArgValCysLeuAlaAsnGlySerTrpSerGly 2881
 DB CTTGAAGTCTCAGACATCCGTAGGACACC-ATGCTGGCTCTCACCCCTCTCTCAGGCAGA 5444
 QY 2882 AlaThrProAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGlyVal 2901
 DB GAGGAGGTGGATGTGTAGATGTGATGTGCTGGGCCAGATGGCTG- 5492
 QY 2902 ThrGluGlyLeuAspTyTyGlyPhe-----MetLysGluValThrPhe- 2915
 DB -----CACCCCATTTGCTTCTCTCCGAGGAGGAGCAGCTCAGATTTGATGATGAAGA 5549
 QY 2916 ---HisCysHisGluGlyTyTyTyLeuHis-----GlyAlaProLysLeuThr 2930
 DB TGAAGATGCAAGAGACTTCTTCTGCTAACATCATCACAGACTTGGTCTACCA-----GGG 5603
 QY 2931 CysGlnSerAspGly-----AsnTrp---AspAlaGluileProLeuCysLys 2945
 DB TGCACCTCCAGCCCGCCAGACAGACCGGACTGTGTGATGATGGCTCGCACCT---TGCAGC 5660
 QY 2946 ProValAsnCysGly-----ProProGluAspLeuAlaHisGlyPhePro 2960
 DB CCGCTACTCAGCGCTGATGTGCTGCCAAGCTCTCT- 5696
 QY 2961 AsnGlyPheSerPheileHisGlyGlyHisileGlnTyTrglnCysPheProGlyTyLys 2980
 DB -----GGATGCGGTGAGATGCAATGC-----CCAGGACAA- 5729
 QY 2981 LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySer---Ser 2999
 DB -----CATGGGCGGTGCTCCACTCCATGCTGC-----AGTGGCAGTGTGATGC 5771
 QY 3000 ProSerCysLeuPro-----CysArgCysSer----- 3008
 DB CCAAGGTGCTTCCAGATCTGTATGCGAACCGAGTAACATGATCTAGATGCCAGGATGAA 5831
 QY 3009 -----ThrProValilleGluTyTyGlyThrValAsnGlyThrAspPheAspCys 3024
 DB TGATGTACTACACCTT-----GATCCT 5855
 QY 3025 GlyLysAlaAlaArgileGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluile 3044

Db 5856 GGC-----TGCCGCGCTGGC----- 5870
Qy 3045 ThrCysGluAlaAspGlycInTrpSerGlyPheProHis---CysGluHisThrSer 3063
Db 5871 ---TGTCGAGGGAATGGTGGCAGAACTGATCAACTGCACGAGCGATGTAA----- 5918
Qy 3064 CysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerTrpLysGlu 3083
Db 5919 TGCACTGGATGACCA-----TGGA--- 5939
Qy 3084 AsnValIleThrTyrSerCysArgSer---GlyTyrValIleGln---GlySer 3099
Db 5940 -----ATCTGCTTTCACCTGGGCGAGCTGCTGTCATAATATGTGGAGCCAC 5984
Qy 3100 SerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluPro 3119
Db 5985 TCTTTTGTGTGT-----GMAAAA-----TGCGGCCAACCG----- 6014
Qy 3120 LeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThr 3139
Db 6015 -----AGACATCGAGGACAA 6029
Qy 3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp 3159
Db 6030 CAGGAGAGACACCTCTGTTCTTCTGCTCCCGGAGGGGAGCTATGAGAGCCACAGAT 5089
Qy 3160 ThrPheThr-----CysGln-LysAspGlyArgTrpPheProGluArgIleSerCysE 3177
Db 6090 CCTGTTAGACCATTTTGCATCGAGACATCACAGACCATATGATGCTG----- 6138
Qy 3177 rProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPh 3197
Db 6139 -----CTTCCCGGAGTGTGCTCGGAGTCGATGACACCATGACATGCT 6182
Qy 3197 eSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAs 3217
Db 6183 GGCCTCTTCGATGATACATGTC----- 6207
Qy 3217 nIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSer-----AspGluSe 3235
Db 6208 -----ACCCCAAGCCCTCCAGGCGCCGCTGTCACCTTCTGTC 6242
Qy 3235 rCysSerProValSerCysGlyLysProGluSerProGluHisGlyPheVal----- 3252
Db 6243 TCTCTCACCTGTCTCTGCG-----CCCAACAGATCTTCTCAGCCCTGAA 6290
Qy 3253 -----ValGlySerLysTyrThrPheGluSerTh 3262
Db 6291 GCACACCCCAATGGGCAAGAGCTAGACGCGCCAGTGCACAGAGTACCCTACTAG 6350
Qy 3262 rIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGl 3282
Db 6351 CCTCCCTAACCTTGCACAGGAGCAAGGATGCAAGGGTAGTAGGAGGAAG----- 6402
Qy 3282 nGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrPr 3302
Db 6403 -----AAGTCTCTGAGTGAGAGGTCCAACTGTCTGAGAGTTCAGTAACCTTATTCCTCC 6455
Qy 3302 oLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrGlyProAsnValVa 3322
Db 6456 TGTGATTCCCTAGATCTCCTCACACGATGTTTCCGACACCCACATCTCTCCATGAT 6515
Qy 3322 lTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAs 3342
Db 6516 TACATCCCTGGGATCTT-ACAGGCGCTCACCAACCTATGTTGGCCACTGCGC----- 6569
Qy 3342 nGlyThrTrpSerHisProValProLeuValLysProAsnProCysPro-----ValPr 3360
Db 6570 -----CCCTCTGCT-----CCAGTCCATGCCCCAGCATGCACTAT 6604
Qy 3360 oPheValIleProGlu---AsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAs 3379

Db 6605 CTTTTTCTAACCTTCATGAATGACAGCCTTTGGC----- 6638
Qy 3379 nValserIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIlelleThrCysAs 3399
Db 6638 ----- 6638
Qy 3399 nProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPr 3419
Db 6639 -----ACATGGGCCAGCAGCTGTGCTTCCCTCAGTGAGCCAGTTGCTATCCACC 6688
Qy 3419 oAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleTh 3439
Db 6689 ACCACAT-----TGTCCTCCAGG-----CA 6709
Qy 3439 rTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGlu-- 3458
Db 6710 GTGGCAGTGTCTGAAG-----CTTGAGTA 6733
Qy 3458 ----- 3458
Db 6734 GGTCTCATCCAGTCCAGTCCAGCAGATGGATGTAACCGCATGGAGTGAATGAGACCC 6793
Qy 3459 -----AsnGlyThrTrpThrSe 3464
Db 6794 AGTACAATGAGATGTTTGTGTATGTTGCTGCTCCAGCTGAGGCGCACCCATCTCTGGCATAG 6853
Qy 3464 rProPro-----IleCysArgal 3470
Db 6854 CTTCCCGAGAGCGGCCACCTGTAAGGGAAGACATACACCCCTCGGGAGCCCTTGGCCCC 6913
Qy 3470 aValCysArgPhePro-----CysGlnAsnGlyGly----- 3480
Db 6914 CCATTGTGACTTTCAGCTCATCTCCCTAAAGGAGATATGGCCCAACAGCGGGGCTCCCC 6973
Qy 3481 ----IleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCy 3499
Db 6974 AGCTCAGTCCACCTGCTCCCTCCAGCTGT-----TG 7003
Qy 3499 sGluGluProIleCysIleLeuProCysLeuAsnGlyGly-----ArgCysValAla 3516
Db 7004 CGGGCCCCCTGCCCCACCATGTACCATTCAGAAATGGCCCCCTTTGGCCCGTGTGGCTT 7063
Qy 3516 aProTyrGlnCysAspCysProGlyTrpThrGlySerArgCysHisThrAlaValCy 3536
Db 7064 TCCCCACTGCCATGATGCCCCAGCAGGAGCGGAGTAGCTCAGACCATTTCTCCAGCCT 7123
Qy 3536 sGlnSer---ProCysLeuAsnGly 3543
Db 7124 ATCATCTTTTCCAGCCTCTGTGGG 7148

RESULT 15

US-09-484-970B-110/c
; Sequence 110, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484, 970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 110
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 009051.SCB1
US-09-484-970B-110

Alignment Scores:
Pred. No.: 4,54e-60 Length: 763
Score: 838.00 Matches: 148
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.20% Indels: 0
DB: Gaps: 0

US-09-977-053-4 (1-3571) x US-09-484-9708-110 (1-763)

```
QY 720 ileValileLysGlySerProCysGluileProPheThrProValaenGlyAspPhele 739
DB 763 ATTGTCATAAAGGCTCTCCCTGTAATTCATTCACACCTGTAATGGGATTTATA 704
QY 740 CysThrProAspThrGlyValaenCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
DB 703 TGCACCTCCAGATAACTGGAGTCAACTGTACATTAACCTTGTGGAGGCTATGATTTC 644
QY 760 ThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTrpLysProThr 779
DB 643 ACGAGAGGCTACTGACAGATTAATTGCTTATGAGATGCGCTCGGAACCAACA 584
QY 780 TyrThrThrGluTrpProAspCysAlaLysLysArgPheAlaenHisGlyPheLysSer 799
DB 583 TATACCACTGAATGGCCAGACTGTGCCAAAAACGTTTTTCAAAACCAACCGGTTCAAGTCC 524
QY 800 PheGluMetPheTyrLysAlaLysArgCysAspAspThrAspLeuMetLysLysPheSer 819
DB 523 TTGAGATGTTTACAAAGCAGCTGCTGTGTGATGACACAGATCTGTGAAGAAGTTTCT 464
QY 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
DB 463 GAAGCAATTTGAGACAGACCTGGGAAAAATGTTCCCATCATTTTGTAGTGATGAGGAC 404
QY 840 ileAspCysArgLeuGluLeuAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
DB 403 ATTGACTGCAGACTGGAGGAGAACCTGACCAAAAAATATTGCTAGAAATAATAATTATGAC 344
QY 860 TyrGluAsnGlyPheAlaLysGly 867
DB 343 TATGAAATGGCTTTGCAATTGGT 320
```

RESULT 16

US-08-939-366-27
Sequence 27, Application US/08939366
Patent No. 6355415
GENERAL INFORMATION:
APPLICANT: Wagner, Thomas E.
APPLICANT: Xie, Yuesheng
TITLE OF INVENTION: Compositions and Methods for the Use of
TITLE OF INVENTION: Ribozymes to Determine Gene Function
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/939,366
FILING DATE: 29-SEP-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: OHU-02749
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 6677 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-939-366-27

Alignment Scores:
Pred. No.: 4,61e-58 Length: 6677
Score: 834.00 Matches: 540
Percent Similarity: 31.42% Conservative: 205
Best Local Similarity: 22.78% Mismatches: 784
Query Match: 4.18% Indels: 848
DB: Gaps: 122

US-09-977-053-4 (1-3571) x US-08-939-366-27 (1-6677)

```
QY 927 ProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeuGlnThrLeu 946
DB 530 CCTGACCCCTGCGAGGATACCCCAACTCTGCAAGAAATGGTGGCAGCTGCAAGCCCTGCTC 589
QY 947 GluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGln 966
DB 590 CCCACACCCCAAGCTCCCGTAGTCTCTCTCCACTGACCCCTCACTTCTCTCTGTCACC 649
QY 967 LeuAlaSerGluLeuLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986
DB 650 TGGCCCTCTGGC---TTACCCGGTGATCGATGCCAAACCCCATCTGGAAGAGCTCTGTCCA 706
QY 987 -----PheCysArgProGlySer-----ValLeuArgGlyArgMetCysVal 1000
DB 707 CTTCTCTTTCTTCCACACGGGGGTCACTGTATGTTTCAGGCTCAGGCCCGCCACAGTGC 766
QY 1001 AsnCysProLeuGlyThrTyrAsnLeuGluHisPheThrCysGluSerCysArgile 1020
DB 767 TCTTGGAGCTGGG-----TGGACAGGTGAGCAATGTCAGCTC 805
QY 1021 GlySerTyrGlnAspGluGlu-----Gly 1028
DB 806 CGAGACTTCTGCTCAGCAACCCCTGTGCAACGAGCGGTGCTGCTGGCCACATACCCC 865
QY 1029 GlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAsn 1048
DB 866 CAGATCCAGTGGCGC---TGTCCACCTGGG---TTCGAGGGTGCACCTGTGGAACGCGAC 919
QY 1049 IleSerAspCys-----LysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065
DB 920 ATCAACAGATGCTCTCTGGAGCCGGGACCTGCTCCTCAGGCGACCTCTCTGCGCATTAAC 979
QY 1066 LeuGluThrCysGluSer---CysProLeuGlyThrTyrGlnProLysPheGlySerArg 1084
DB 980 TTGGGTTCCTACCACTGTCTCTGCTCTGGGCGAGGAAGTCCCGCAGCTCAGG 1039
QY 1085 -----SerCysLeuSer-----CysProGlu 1091
DB 1040 AAGGGAGCTTGGCTCTCTGGAAGCTGTCTCAATGGGGGGCAGCTTCCAGCTGCTCCAGAG 1099
QY 1092 AsnThrSerThrValLysArgGlyAlaValaAsnIleSerAlaCysGlyValPro----- 1109
DB 1100 GGACACTCCACC-----TTTCATCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 1138
QY 1110 -----CysProGlu 1112
DB 1139 TTCACGGGGCTGGACTGTGAGATGAACCCAGATGACTGTGTGTGTCAGGCGACCACTGTCA 1198
QY 1113 GlyLysPheSerArgSerGlyLeu-----MetProCysHisProCysProArgAspTyr 1130
DB 1199 GGGGCCACCTGTCTGGATGGGCTGGATACCTACACCTGC-----CCCTGCCCCCAAGACATGG 1255
```

QY 1130 ----- 1130
Db 1256 AAGGCTGGAGTCTCTGAAGATATAGATGAATGTGAAGCCGGGGTCCCCCTCGCTGC 1315
QY 1131 ----- TyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPhe 1145
Db 1316 AGGAACGGTGGACCTGCCAGAACACAGCTGGCAGCTTTCACCTGTGTGGTGGTGGTGGC 1375
QY 1146 TyrGly----- ThrThrProPheAla 1152
Db 1376 TGGGAGGTGGCAGGCTGTGAGGAGAACCTGGATGACTGTGCAGCTGCCACCTGTGCCCG 1435
QY 1153 GlySerArgSerIleThrGluCysSerPheSerThrPheSerAlaAlaGluGlu 1172
Db 1436 GGATCCACCTGCATCGACCGTGTGGCTCTTCTCC----- 1471
QY 1173 SerValValProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGln 1192
Db 1472 TGCCTCTGCCACCTGGACGCACAGGCTCTCTGTGCCACCTGGAAGACATG----- 1522
QY 1193 ValPheHisGluCysPhePheAsnProCysHisAsnSerGlyThrCys----- GlnGln 1210
Db 1523 ----- TGTTCAGTCAGCGCTGCCACGCTGAATGCCAGTGCAGCACCAACCCCT 1570
QY 1211 LeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThr 1230
Db 1571 CTGACAGGCTCCACCTCTGCATATATGCCAGCTGGCTACTCAGGATCCACCTGTGCACCA 1630
QY 1231 AsnIleAspGluCys----- SerProLeuProCysLeuAsnAsnGlyVal 1245
Db 1631 GATCTGGATGAGTGCMAATGGCCAGCAAGACCCAGTCCCTGCGAATCGGGGGTCC 1690
QY 1246 CysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGlnArg 1265
Db 1691 TGCATCAACACCCCTGGCTCTTCACTGCCTCTGCCTGCCTGGTGTACACGGGCTCCCGC 1750
QY 1266 CysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIleCysVal 1285
Db 1751 TGTGAAGCTGACCAATAGTGTGCTCTCACTGCCTCTGCCTGCCTGGTGTACACGGGCTCCCGC 1810
QY 1286 AspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHisCysGlu 1305
Db 1811 GACCTGTTCACCTTCCACTGCCTCTGCTCCACCCAGGCTTGGAGGGGAGACTCTGTGAG 1870
QY 1306 ThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGluAspGln 1325
Db 1871 GTGGAGTCAATGAGTGCACCTCTAATCCCTGCCTGAAACCAAGCTGCCTGCCATGACCTG 1930
QY 1326 ValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGlyLysAsn 1345
Db 1931 CTCAACGGCTTCCAGTGCCTCTGCTTCTGATTCACGGGCGCCGATGTGAGAAAGAC 1990
QY 1346 ValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGlyAlaAsn 1365
Db 1991 ATGGACGAGTGTAGCAGCACCCCTCTGCCATGGGGGGCTGCGGACCCAGCTGGA 2050
QY 1366 SerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsnIleAsn 1385
Db 2051 GCCTTCTACTGCGAGTGTCTCCAGGCTTTGAAGGGGCCACACTGTGAGAAAGAGTGGAC 2110
QY 1386 GluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsnSerTyr 1405
Db 2111 GAATGCTGAGTGACCCCTGTCCGCTGGAGGAGCAGCTGTCTGTGATCTCCCGGAGCATTC 2170
QY 1406 SerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSerThr--- 1424
Db 2171 TTCTGTCTGCGCTCTGCTTTTACAGGTCAACTTTGTGAGGTTCCCTGTGTGCACCCGCC 2230
QY 1425 ----- GlyPheAsnLeuAspPheGlu----- 1431
Db 2231 AACATGTGCCAACCTGGACGACGAATGCCAAGGTCCAGGAACACACAGAGCCCTGCTTGC 2290
QY 1432 ValSerGlyIleTyrTyrValMetLeuAspGlyMetLeuProSerLeuHis-AlaLe 1451

Db 2291 CTTGACGGAAGTCTCTGGCTGTCTCTGCGGAGGACAACTGCCCCCTGTCACCATGGCCAT 2350
QY 1451 uThr----- CysThrPheThrMetLysSerSerAspAspMetAsnTyrGlyThrPr 1468
Db 2351 TGCCAGAGATCTCTGT 2399
QY 1468 oileSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeuThrAspTyrAsnG1 1488
Db 2400 ----- GCGAGACAGAAC----- 2411
QY 1488 YTriPVal-LeuTyrValAsnGlyArgGluLysIleThrAsnCysProSerValAsnAspG 1508
Db 2412 -TGGTGGCTGCATCTCCAC----- ACCCTGTGCCCA----- 2442
QY 1508 LysArgTriHisIleAlaIleThrThrSerAlaAsnGlyIleTTP----- LysValT 1527
Db 2443 ----- TGGGGGACCTGCCACCCACA----- GCCATCTGGCTACAACTGT 2482
QY 1527 YrIleAspGlyLysLeuSerAsp----- GlyGlyAlaGlyLeuSerVal- 1541
Db 2483 ACCTGCCCTGCAGCTTACATGGGGTTGACCTGTAGTGAGAGGTGACAGCTTGTACTCA 2542
QY 1542 ----- GlyLeu----- 1543
Db 2543 GGGCCCTGTCTCAATGGTGGCTCTCTGCAGCATCCGTCTCTGAGGGCTATTCTCTGCCACTGC 2602
QY 1544 ----- ProIleProGly----- 1547
Db 2603 CTTCCAAGTCAACAGAGTGCCTGCCACTGCCAGACTGCCGTGGACCACTGTGTGTCTGCTCG 2662
QY 1547 ----- 1547
Db 2663 TGCCTCAATGGGGGTACCTGTGTGAACAGCTGGCACTTCTTCTGCTCTGTGTGCCACT 2722
QY 1548 ----- GlyGlyAlaLeuValLeuGlyGlnGlu----- GlnAspLysL 1560
Db 2723 GCTTCCAGGGGCTGCATCTGTGAGGAGAGACTAACCCTCCAGCTGTGCAGACAGCCCTGC 2782
QY 1560 yGlyGlyGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuT 1580
Db 2783 AGGAACAAGGCAACCTGCCAAGACACACTCGAGGGGGCCGCTG----- CCTCTGACGCCCT 2839
QY 1580 rAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluL 1600
Db 2840 GCTATATC----- AGAAGACAGCTGCCAGACTCTGATAGACTGTGTGCCCCGGAAGCCC 2893
QY 1600 euserLysGlyAsnValLeu----- AlaTriP-ProAspPhe----- LeuSer 1613
Db 2894 TGTCCACACACTGCTCGATGCCTCCAGAGTGGGCGCTCTGTCAGTGCCTCTCCAG 2953
QY 1614 GlyIleValGlyLysVal----- LysIleAspSerLys 1624
Db 2954 GGATGACAGAGGGGCTCTCTGTGACTTCCACTGTCTGTCAGAACGCCCGGATGAGCCAA 3013
QY 1625 SerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAla 1644
Db 3014 GGCATAGAGATCTCTGGCTGTGCCAGATGAGGGCTCTGT----- 3055
QY 1645 SerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln--- 1663
Db 3056 ----- ATTGACACGGGCTCTCTATTTCTGCGCTGCGCTCTCTGATTTCCAAGGC 3106
QY 1664 ---LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnThrThrGlnProLeu 1682
Db 3107 AAGTTATGCCAGGATTAATGTGAACCCCTGC----- GAGCCCAAT 3145
QY 1683 ProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHisSer 1702
Db 3146 CCC----- TGCCATCAGGGTCTACTCTGTGTGCTCTCAGCCC----- 3181
QY 1703 AlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeu 1722

QY 2339 hrThrGluVal-----GlyValValThrPheSer---CysLysGluGlyH 2353
DB 4878 ACCAGAGCCGCAAGTGGATGAGGTGAGTGCCTGCGCCCTGAGAGGAG 4937
QY 2353 isValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerP 2373
DB 4938 AGGCTGAAGAAACAGCCTCAGCCTCAGGTGC-----CAGCTTTGG----- 4978
QY 2373 heProValCysLysIleValLeuCysThrProProLeuIleSerPheGlyValProI 2393
DB 4979 -----CAGCTCCTCAGCCTCAGGTGC-----CAGCTTTGG----- 4978
QY 2393 leProSerSerAlaLeuHisPheGlySerThrValLysThrCysValGlyGlyPheP 2413
DB 5004 TCCCCCAGGACCCATG-----CAGCTCCTCAGCCTCAGGTGC----- 4978
QY 2413 heLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeuP 2433
DB 5020 -----CAGCTCCTCAGCCTCAGGTGC-----CAGCTTTGG----- 4978
QY 2433 roGluCysValProValGluCysProGlnProGluGluLeuProAsnGlyIleIleAspV 2453
DB 5021 -----CAGCTCCTCAGCCTCAGGTGC-----CAGCTTTGG----- 4978
QY 2453 alGlnGlyLeuAlaThrLeuSerThrAlaLeuThrCysLysProGlyPheGluLeuV 2473
DB 5061 TGGAC-----ACCTGTGGACCTGATGGGGTGGACAC 5090
QY 2473 alGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrC 2493
DB 5091 CCTGTATGCTCAGCGCTCTCTGTGGGGAGTGCAGTCCACGACTGGGGTAGTCCACAGA 5150
QY 2493 ysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerThrA 2513
DB 5151 GACTGGGGCTAGGAAATCTGGAACTCTGGAACTCTG----- 5188
QY 2513 spLeuHisThrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyP 2533
DB 5189 -----CTGGATAGAGG-----GCTGCC 5207
QY 2533 roSerAlaLeuThrCysLeuGluThrGlyAspTrp---AspValAspAlaProSerCys- 2551
DB 5208 CCCAGGCTCA-CACGT-----GGGCACTGGAGAGCGCTCTGCACCTAGCTGCC 5257
QY 2552 -----AsnAlaIleHisCys-----AspSerProGlnProI 2562
DB 5258 AGATTCTCTCGGCCCAACCGCTGCGCGCGCTCTTGGAGCTGGAGCCCAACCCCAACAG 5317
QY 2562 leGluAsnGlyPheValGluGlyAlaAspThrSerThrGlyAlaIleIleThrSerC 2582
DB 5318 CCAGACCGCGCTGGCGCGCACCCCACTTCACTGCTGTGGC----- 5358
QY 2582 ysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrps 2602
DB 5359 -----TGGCGACGCTCGGCA---G 5374
QY 2602 erSerSerIleProThrCysMetProIleAspCysGlyLeuProProHis----- 2618
DB 5375 GTTTGCCAGCTCTATTATGGCCAGCAGACAGACTACGGTGGAGCGCCCGCACAGAGACGG 5434
QY 2619 -----IleAspPheGlyAspCysThrLysLeuLysAspAspGlnGly---TyrPheG 2635
DB 5435 ACTACACCTTTGATGCTGCG---TGCCAGGCTGGCGCTGGAGGACCTGGTGAAGATTG 5491
QY 2635 luGlnGluAspAspMetMetGluValProTrpVal----- 2646
DB 5492 ATCG-CAGCCCGCAGCAGATGAGGAGCCAGGCGATAAAGGGGAAAAATCGCACTGCACCTG 5550
QY 2647 -----ThrProHisProProThrHisLeuGlyAlaValAlaLys----- 2659
DB 5551 GCGCGCTGCTGGAACAACCGCGAGCGCGCTCTCTCTCCAGGCTGGAGCGGATAA 5610
QY 2660 -----ThrTrpGluAsnThrLysGluSerP 2668

DB 5611 AGATGCCAGGACACTGAGGACACAGCGCGCTTCTTGGCAGCGCGGAGGAGCGGT 5670
QY 2668 roAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerThrCysAsnProG 2688
DB 5671 CGAGGTGGCGCAGCTGTGCTGGAGCTCGGGCGCGCCCGGAGACTCGCAGACCAGCGCG 5730
QY 2688 lyThrGluLeuLeuGlyAsnProValLeuIleCysGlnGlu----- 2701
DB 5731 GCTGCGCC-----CAGGAGATGTGGCGCGCCGCGCAGCGCAG 5763
QY 2702 --AspGlyThr-----TrpAsnGlySerAlaProSerCysIleSerIleGluC 2717
DB 5764 TCACCTGGGACCTCTTAACGCTCTGGAA-GGGGCTGGACCGACTACGACGAGGCGCGTG 5822
QY 2717 ysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlyS 2737
DB 5823 CGCAGCGACGACACCGCGCGGGGCG-----GGGT 5852
QY 2737 erAlaValGlnThrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg--- 2755
DB 5853 CGGCCCCCGCG-----TGCAGGACG-----CTGTCTCGGGAGCGCGCCCGCGGGG 5900
QY 2756 -----LeuCysLeuGluAsnArgLysTrpSer-----GlyAlaSer----- 2767
DB 5901 GCGGAGCTGTCTGCGAGCTCGCCTTGGTGGTGGAGCGCGCGGAGGAGAGG 5960
QY 2768 -----ProArgCysGluAlaIle-----SerCysLysLysProAsnProValMetAsnG 2784
DB 5961 TGTATGCTCGCTGCGGAGCGCGATCTGGAAGCTCGGG-AGGCCCCACACCGCGCGCGCG 6019
QY 2784 lySerIleLysGlySerAsnThr-----TyrLeuS 2795
DB 6020 AGTTCTCGCGGGCTGCGGAGCGCGCGGGCTAGGGCATCACAGATGACTGGCCT 6079
QY 2795 erThrLeuTyrThrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgThrC 2815
DB 6080 CGCGACTG-----GGTGGCCCTGGAAGCTG-----CGGCTCGCGCTGCAGT 6121
QY 2815 ysGlnAspAspLysAsnTrpAspGluAspGluPro---IleCysIleProValAspCys- 2833
DB 6122 GCGCGATCCCGCTCCAGCGCTGAGCCCGCTGCCAGAACGCTGGATCCCTCAAGTTGCC 6181
QY 2834 -----SerSerPro-ProValSerAlaAsnGlyGlnValArg----- 2845
DB 6182 TGGGCTTTCAGATTCCACAGAGATTCCCTTAAGCTGGTTGTAGAAATCTGAAGTAG 6241
QY 2846 -----GlyAspGluTyr-----ThrPheGln 2852
DB 6242 GCAGCTGGGTGGAAGGAGGAGCGACGCTAGGAGTCTGGAAGACTCCGCGACTTTTAAG 6301
QY 2853 LysGluIle-----GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSer 2870
DB 6302 GCCAAATAACCGTTAAGCTCACTTGT-----CTCCCCCAT 6337
QY 2871 ArgValCysLeuAlaAsnGly 2877
DB 6338 AGAGTATGACAGCAATGGA 6358

RESULT 17

US-09-467-997-6
; Sequence 6, Application US/09467997
; Patent No. 6379925
; GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uyttendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6

```

; LENGTH: 6677
; TYPE: DNA
; ORGANISM: mouse
US-09-467-997-6

Alignment Scores:
Pred. No.: 4,61e-58 Length: 6677
Score: 834.00 Matches: 539
Percent Similarity: 31.24% Conservative: 203
Best Local Similarity: 22.69% Mismatches: 784
Query Match: 4.18% Indels: 854
DB: 4 Gaps: 119

US-09-977-053-4 (1-3571) x US-09-467-997-6 (1-6677)

Qy 927 ProAspGluArgAsnAspThrLeuGluTrpGluAenGlnArgLeuLeuGlnThrLeu 946
    ||||| ||||| :|||
Db 530 CTTgACCCCTGCAGGGATACCAACTCTGCAGAAGATGTGTGCAGCTGCCAAGCCCTGCTC 589

Qy 947 GluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGln 966
    ||||| :|||
Db 590 CCCACACCCCAAGCTCCCGTAGCTCTACTCTCCACTGACCCCTCACTTCTCTCGTCACCC 649

Qy 967 LeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerPro 986
    ||||| |||
Db 650 TGCCCCCTCTGC---TTCACCGGTGATCGAATGCCAAACCCATCTGGAAGAGCTCTGTCCA 706

Qy 987 -----PheCysArgProGlySer-----ValLeuArgGlyArgMetCysVal 1000
    ||||| |||||
Db 707 CCTTCTTTCTTCCACAGGGGGTCACTGTATGTTTCAGGCCTCAGGCCGCCACAGTGC 766

Qy 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
    ||||| |||||
Db 767 TCTTCGAGCTGGG-----TGCACAGGTGAGCAATGCCAGCTC 805

Qy 1021 GlySerTyrGlnAspGluGlu-----Gly 1028
    :|||
Db 806 CGAGACTTCTGCTCAGCCAAACCCCTGTGCCAACGGAGCGGTGTGCTGGCCACATACCCC 865

Qy 1029 GlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAsn 1048
    ||||| ||||| |||||
Db 866 CAGATCCAGTCCCGC---TGTCACCTGGG---TTCAGGGGTACACCTGTGAACGGAC 919

Qy 1049 IleSerAspCys-----LysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGly 1065
    ||||| ||||| |||||
Db 920 ATCAACGAGTCTTCTCTGGAGCCGGGACCTGCCCTCAGGGCACCTCTGCCATAACACC 979

Qy 1066 LeuGluThrCysGluSer---CysProLeuGlyThrTyrGlnProLysPheGlySerArg 1084
    ||||| :|||
Db 980 TTGGGTTCCTACCAAGTGTCTGTGCGCTGTGGGGCAGGAAGTCCCCAGTGCAGCTCAGG 1039

Qy 1085 -----SerCysLeuSer-----CysProGlu 1091
    ||||| |||||
Db 1040 AAGGGACCTGCCCTCTCTGGAGCTGTCTCAATGGGGGCACCTGCCAGCTGGTCCACAG 1099

Qy 1092 AsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValPro----- 1109
    ||||| |||||
Db 1100 GGACACTCCACC-----TTTCATCTCTGCTCTGTCTGCCCAGGT 1138

Qy 1110 -----CysProGlu 1112
    |||||
Db 1139 TTCACGGGGCTGGACTGTGAGATGAACCCAGATGACTGTGTGACGGCACCAAGTGTGCAAC 1198

Qy 1113 GlyLysPheSerArgSerGlyLeu-----MetProCysHisProCysProArgAspTyr 1130
    ||||| |||||
Db 1199 GGGGCCACACTGTCTGTGATGGGCTGGATACCTACACCTGC---CCCTGCCCAAGACATGG 1255

Qy 1130 ----- 1130
Db 1256 AAGGGCTGGGACTGCTCTGGAAGATATAGATGAATGTGAAGCCCGGGGTCCCCCTCGGTGC 1315

Qy 1131 -----TyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPhe 1145
    ||||| |||||

```


[illegible]

Qy	1736	Ttp-----AenGlyValSerProSerCysLeu-----	174
Db	3281	TGTACTCCAGCGCTGGAGGCTTCCACTGTGCTCTGCCCTCCAGGCTTCGTGGGACTGCGC	3340
Qy	1745	-----AspValAspGluCysAlaValGlySerAspCys-----SerGluHisAla	1759
Db	3341	TGTGAGGAGATGGATGAGTGT-----CTGGACGGCCCTGTACCCTCGGCACACTGCA	3397
Qy	1760	SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp	1779
Db	3398	GCTTGCCACTCTTAGCCAAAGCCTTCTACTGCCAGTGTCTGCTGGGACACACA-----	3451
Qy	1780	GlyLysAsnCys---AlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGly	1798
Db	3452	GGCCAGCGGTGTGAGGTGAGATGAGCTCTGTGAGGC-----CAACCCCTGCTCCAAT	3505
Qy	1799	HisSerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGly	1818
Db	3506	GGAGGATCTGTGAGATC-----ACAACAGGCGCCACCCCTGGCTTTCACCTGTAC	3556
Qy	1819	TyrGlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeu	1838
Db	3557	-----TGCCCAAGGGTTTGAAGC-----	3577
Qy	1839	IleProTyrCys-----LysAlaValSerCysGlyLysProIleLeuProGluAenGly	1856
Db	3578	---CCCACTGCAGCCACAAAGCCCTTCTCGGGCAT-CCATCACTGCCACAATGGAGG	3633
Qy	1856	Y---CysIleGluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys	1875
Db	3634	CCATATGCTGCCC-----TCCCTTAAGCCAGGGTCAACACCGCTGTGCTGCCTCGAC	3687
Qy	1875	sGlyTyrThrLeuAlaGlyAspLysGlu-----	1884
Db	3688	TGGTTT-----GGGGCCCTGACTGTCTGCACCTCCAGCTCCACCGGGCTGGG	3738
Qy	1885	-----SerSerCysLeuAlaAsnSerTrpSerHisSerProProValCysGluPr	1902
Db	3739	TCCCCCTCACCTGCCCTGCACAATGTACTGTCACTGCAGCCCTGGTGTGGGCAACCC	3798
Qy	1902	o---ValLysCysSerSerProGluAsnIleAsnAenGlyLysTyrIleLeuSerGlyLe	1921
Db	3799	GGGCTTTCATGACACCTGCCCTCTGACTCTCCAGGCGCCCGGTGTCAAGGCCAGGGGC	3858
Qy	1921	uThrTyrLeuSer-----ThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnI	1939
Db	3859	AGTGGGTGTGGGGCGGAGGTGGTGTATGGCACTTCGGATGCTGGCTGCAGT-----GG	3912
Qy	1939	yProSerIleIleGluCysThrAlaSerGlyIleTrpAsp-----	1952
Db	3913	CCCA-----GGAGGAGACTGGGATGGAGGGGACTGTTCCTCTGGG	3951
Qy	1953	-----ArgAlaProPro-----AlaCysHisLeuValPheCysGly	1964
Db	3952	GGTCCAGACCCCTGGAAGGGCTGTCCCCCGGCATTCCTCAGTGTGGCTTCTGTTC	4006
Qy	1964	yGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsnTh	1984
Db	4006	-----	4006
Qy	1984	rValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCysLe	2004
Db	4007	-----CGGAGCGAGCGGTCTCACCGCGAGTGTACTCTGAGGAGTGTCT	4050
Qy	2004	uAlaAspGlyLysTrpSerArgSerAspGlnCysLeuAlaValSerCysAspGluPr	2024
Db	4051	CTTTGATGGC-----TACGACTGTGAATCCC	4077
Qy	2024	oProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhePheGlyAspIleAlaPh	2044
Db	4078	TCCAACTGCATCCAGCCTATGACCG-----	4105
Qy	2044	eTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGly	2064

Db 4106 ----TACTCCGAGATCACTTCCACACCGGCACTGTGAGAAAGGCTGCAATACGCTGA 4161
QY nGlyLysTrp-----ValProProGluGlyGlnAsp-----MetProAr 2077
Db 4162 ATGTGGCTGGAGCGGGAGAGACTGCAGACACAGAGGGGAGAGACTCAGAGGGGAGGCGCTC 4221
QY 2077 gCysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerVa 2097
Db 4222 CTTGGCCCTGTGGTGGTGTGAGGGCGCCCGAGCCCTGGATCAGAGAGCTGCTT----- 4273
QY 2097 lSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheVa 2117
Db 4274 -----GCCCTGCACAGTGTGTCTCTG----- 4297
QY 2117 lleuAenThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMe 2137
Db 4298 -----ACTCTGAGGGTCTGGTCTCTG----- 4318
QY 2137 tSerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAl 2157
Db 4319 -----GTGAGGAGGACAG 4332
QY 2157 aSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTy 2177
Db 4333 TGAAGGACAGACATGTGTTC-----CCCTATCTCTGGACCGGGCC----- 4375
QY 2177 rIleLysGlyGluLysSerThrCysGluAlaThrGlyGlnTrpSerProIlePr 2197
Db 4376 -----AAAGAGGAGTGTAGTGGAGTAGGATTCCTCTTCATGGGAAAGACAAAGCCCC 4428
QY 2197 oThrCysHisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHi 2217
Db 4429 TCCACTCAGCCCTGGGC---AAGGAGACAGAGTCTCTTGGTGGAGGGTTGTG----- 4480
QY 2217 sThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSe 2237
Db 4480 ----- 4480
QY 2237 rValGlySerProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLe 2257
Db 4481 -----GTAGTGATGGAGTGGATCTGTCCCGCTG-TGGTCCGGAAACATCC----- 4524
QY 2257 uMetCysValProLeuAspCysGlyLysProProIleGluAsnGlyPheMetLysGl 2277
Db 4525 ----TGGTCTCCCTGCCCTGGGACTCTGGACTCTCTGCTGCGCTTCTTGGACGCAATGG 4580
QY 2277 yGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuVa 2297
Db 4581 CAGCAGTGGAGCTCTGGAGCCCTGTGCTGGACCTTGTGCGGGCTCACCCCTCAAG 4640
QY 2297 lGlyAspSerSerTrpThrCysGlnLysSer----- 2307
Db 4641 CAGGAGACAGGCC---CCCTGCCAACAGCTTCCCTGGCCCATTTCTATGTTACCAAGTGG 4697
QY 2308 -----GlyLysTrpAsnLysLysSerAsnProLysCysMet----- 2319
Db 4698 TTGGGGTGTCTTCTCTGGCCCTTGGGGCCCTTCTGCTCTCCAGCTCATTTGGCGGAGCGG 4757
QY 2320 -----ProAlaLysCysPro----- 2324
Db 4758 GACGAGAACATGGGGCCCTGTGGCTGCCCTGTTTCATTGGAAGGCTCAGACACAGC 4817
QY 2325 -----Glu-ProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuTr 2339
Db 4818 AGGCACCCACCGCGGAGGCCCCCACTTGGCGGAGGACAACTTGTGCTTTAAGGCACTGA 4877
QY 2339 hThrGluVal-----GlyValValPheSer---CysLysGluGlyH 2353
Db 4878 AGCCAGAGGCCGGAAGTGAAGATGGAGTGGAGTGGCCATGTGCTCGGGCCCTGGAAGGGAG 4937
QY 2353 iValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerP 2373

Db 4938 AGGCTGAAGAAACACAGCCCTCAGCCCTCCAGGTGC-----CAGCTTTGG----- 4978
QY 2373 heProValCysLysLysIleValLeuCysThrProProLeuIleSerPheGlyValProI 2393
Db 4979 -----CCGCTCAACAGCGCGCTGTGTGAGAGC 5003
QY 2393 leProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGlyPheP 2413
Db 5004 TCCCCAGGAGCAGC----- 5020
QY 2413 heLeuArgGlyAsnSerThrLeuCysGlnProAspGlyThrTrpSerSerProLeuP 2433
Db 5020 ----- 5020
QY 2433 roGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIleIleAspV 2453
Db 5021 -----CTGACCCCTCTCAGGAGTGTGAATCGGAGGTCTTGGATG 5060
QY 2453 alGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuV 2473
Db 5061 TGGAC-----ACCTGTGGACCTGATGGGTGACAC 5090
QY 2473 alGlyAenThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyLysProThrC 2493
Db 5091 CCCTGATGTTCAGCGCTCTCTGTGGGGAGTGCAGTCCAGCTGGGGCTAGTCCACAGA 5150
QY 2493 yLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThra 2513
Db 5151 GACTGGGGCTAGGAATCTGGAACTCCGGAACCACTG----- 5188
QY 2513 sPLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyP 2533
Db 5189 -----CTGGATAGAGG-----GCCTGCC 5207
QY 2533 roSerAlaLeuThrCysLeuGluThrGlyAspTrp---AspValAspAlaProSerCys- 2551
Db 5208 CACAGGCTCA-CACTGT-----GGCAGCTGGAGAGAGCCCTCTGCTAGCTAGCTGCC 5257
QY 2552 -----AnAlaIleHisCys-----AspSerProGlnProI 2562
Db 5258 AGATTCTCTGGCCCAACCGCTGCCCGCTCTCTGAGGCTGGAGCCCAACCCACAG 5317
QY 2562 leGluAenGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleTyrSerC 2582
Db 5318 CCAGACCGGCTGGCGCACCCCACTTCACACTGCTGTGGC----- 5358
QY 2582 yPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrps 2602
Db 5359 -----TGCCGAGCTCGGACCTGGAGGACCTGTTTGAAGAATTG 5374
QY 2602 erSerSerIleProThrCysMetProIleAspCysGlyLeuProProHis----- 2618
Db 5375 GTTTGCAGCTCTCTATTGGCCAGACAGACTACGCTGGAGCGCCGAGAGAGAGCGG 5434
QY 2619 -----IleAspPheGlyAspCysThrLysLeuLysAspAspGlnGly---TyrPheG 2635
Db 5435 ACTACACCTTTGATGCTGGC---TGCCAGGCTGGCGCTGGAGGACCTGTTTGAAGAATTG 5491
QY 2635 luGlnGluAspAspMetMetGluValProTyrVal----- 2646
Db 5492 ATCG-CAGCCGAGCAGATGTAGCAGCCAGGGATAAAGGGGAAAACTGCACTGCAC 5550
QY 2647 -----ThrProHisProProTyrHisLeuGlyAlaValAlaLys----- 2659
Db 5551 GGCGCTGTGTGTGAACAAACCCCGAGCCCGCTCTCTCTCTCCAGGCTGGAGCGGTAA 5610
QY 2660 -----ThrTrpGluAsnThrLysGluSerP 2668
Db 5611 AGATGCCAGGACAGTAGGAAACAGACCGCGCTTTTCTTGGCAGCGCGCGAGGAGCGGT 5670
QY 2668 roAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProG 2686
Db 5671 GGAGGTGGCGCAGCTGTGTGTGAGCTCGGGGCGCGGCGGAGCTGGAGAGACGAGCGCG 5730

1231 ATCTGTAATGAGCCCTGGCGAGTCTCTGTGAGACCAACTGGGGCGCCAGCTCTGT--- 1287
QY SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
1288 -----GACAAAGATCTCAATTACTGTGGG 1311
QY 1108 val-----ProCysProGluGly-----LysPheSer 1116
1312 ACTCATCGCGGTCTCAACGGGGGAACTGTAGCAACACAGCCCTGACAAATATACG 1371
QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
1372 TGTTC-----TGCCCTGAGGGGTATTCAGACCACTCT--- 1407
QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
1407 ----- 1407
QY 1157 IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluSerValPro 1176
1407 ----- 1407
QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
1408 -----GAAATTCCTGAG-----CAGCC 1425
QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
1426 TGCCTCTCTGATCCCTGTGTCAACAGAGGCGAGCTGTAAAGGAGACCTCCCTGGGCTTTGAG 1485
QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
1486 TGTGAGTGTTCCTGAGCTGGAGCGGCCCCACATGCTCTCAAAACATGATGACTGTCT 1545
QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
1546 CCTAATAACTGTTCCTCCGCGGGGACCTGCGAGACCTGTGTAACGANTTAAGTGTG 1605
QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1276
1606 TGCCCCCAGAGTGGACTGGGAAACCTGCCAGTGTAGATCAAAATGAATGTGAGGCCAAA 1665
QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
1666 CCTTGTGTAAACGCCAAATCTCTGTAGAAATCTCATTTGCCAGCTACTACTCGGACTGTCT 1725
QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
1726 CCCGCTGGATGGTTCGAATTTGTGATTAATTAATTAATGATCTGCTTGGCCAG---TGT 1782
QY 1317 LeuAsnAlaValCysGluAspGlnValGlyPheLeuCysLysCysProProGly 1336
1783 CAGAAATGAGCCCTCTCTGCGGATTTGGTTAATGTTATCGCTGTATCTGTCCACTGGC 1842
QY 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
1843 TATGAGGGCGACTCCTGTGAGAGAGACATCATGAAATGTGCCAGCAACCCCTCTTTGAAT 1902
QY 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
1903 GGGGGTCACTGTGAGAAATGAATCAACAGATTCAGTGTCTGTGTCCACTGTTTCTCT 1962
QY 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
1963 GGAACCTCTGTGAGCTGGACATCGATTATTGTGAGCCTAATCCCTGCCAGAACGGTGC 2022
QY 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
2023 CAGTGTCTAACACCGTGCAGTGAATTTCTGCAAGTGCCTCCGAGGAGCTATGAGGCGCAG 2082
1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436

2083 AACTGTCTCACACCTGAAAGACCACCTGCCGCGACGACCCCTGTGAAGTG----- 2130
QY GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
2131 -----ATTGAC-----AGCTGCACAGTGCC 2151
QY 1457 MetLysSerSerAspAspMetAsnTyrThrPro-----IleSerTyrAlaValAsp 1474
2152 ATGCTTCCACAC-----ACACCTGAAGGGTGCCTATATTTCTCTCC 2196
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
2197 AAC-----GTCGTGGTCTCTCAC 2214
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAla 1514
2215 GGAAG-----TGCAAGATCAGTCGGGNGGCAAAATTC----- 2247
QY 1515 IleThrTyrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
2248 -----ACCTGTGACTGTAAACAAAGCTTCACGGGAACATCTGCCCATGAATAATTAATGAC 2304
QY 1535 GlyAlaAlaGlyLeuSerValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGly 1554
2305 TGTGAGAGCAAC-----CCTTGTAGAAACGGTGGCCTTGCATCGATGGT 2349
QY 1555 GlnGluAsnLysLysGlyGluGlyPheSerProAlaGluSerPheValGlySerIle 1574
2350 GTCAACTCTCTACAG-----TGCACTGTGTAGTACGGCTGGAGGGCCCTAC 2397
QY 1575 SerGlnLeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
2398 TGTGAACCAATTAATTAATGATCAGCAGCAACCC----- 2433
QY 1595 SerCysProGluGluSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
2434 ---TGCCACAAT----- 2442
QY 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
2443 -----GGGGGCGACGTTCGCGACCTGGTCAATGACTTCTACTGTGACTGTAAAAATGG 2496
QY 1634 LeuGlyGlySerValProHisLeuAlaThrAlaSerGluAspLeuLysProGlySerLys 1653
2497 TGAAGAAGAAAGACCTCCACTCAGTGACAGTCAGTGTGATGAGGCCACCTGCAACAC 2556
QY 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
2557 GGTGGCACCTGCTATGATGAGGGGAGTCTTTTAAG----- 2592
QY 1672 CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
2593 TGCATGTCTCTGGCGCTGGGAAGAACCTGTAACTAGCCCGAACAGTAGTGTGC 2652
QY 1692 ValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThr 1711
2653 CTGCCCCAACCCCTGCCATATGG----- 2676
QY 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
2677 -----GGCACATGTGTGTCAACGGCGAGTCTCTTACGTGGCTC 2715
QY 1732 AspAsnGlySerTyrPheGlnValSerProSerCysLeuAspValAspGluCysAlaVal 1751
2716 TGCAGAAGAGCTGGGAGGG-----CCCATCTGTGCTCAGAAATACC----- 2757
QY 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
2758 ---AATGACTGAGCCCTCATCTCTGTTACACAGCGGACCTGTGTGGATGGAGACAC 2814
QY 1767 SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779
2815 TGTGACCGGTGCGAATGTGCCCCGGGTTTGTGTGGCCCGACTGCAGATAAATCAATCAAT 2874

QY 1780 -----GlyLysAsnCysAlaGluProIle----- 1787
DB 2875 GAATGCCAGTCTTACCTTGTGCTTTGGAGGACCTGTGTGATGAGATCAATGGCTAC 2934
QY 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
DB 2935 CGGTGTCTGCTGCTT-----CCAGGCGACAGTGGTGGCCCAAG----- 2970
QY 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrIle 1827
DB 2971 -----TGCCAGGAA-----GTTTCAGGAGGA 2991
QY 1828 ThrCysLeuGluSerGlyGluThrPheHisLeuIlePro----- 1840
DB 2992 CCTTGATCACCATGGGAGT-----GTGATACAGATGGGCGCAAAATGGGATGAT 3042
QY 1841 -----TyrCysLysAlaValSerCysGly 1848
DB 3043 GACTGTATATACCTGCCAGTGCCTGTAATGACGGATCGCTCTCAAGGTCTGGTGTGC 3102
QY 1849 ---LysProAlaIleProGluAsnGly-----CysIle 1858
DB 3103 CTTGACCTTGTCTGCTCCACAAGGGCACAGGAGTGGCCCGGCGAGCTGCATC 3162
QY 1859 -----GluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys 1875
DB 3163 CCCATCCTCGGACGACAGTGCTTC-----GTCACCCCTGC----- 3198
QY 1876 GlyTyrThrLeuAlaGlyAspLysGluSerCysLeu----- 1888
DB 3199 -----ACTGGTGGCGGAGTGTGCTTTCAGTCTCCAGCGGTGAAGACAAGTGC 3252
QY 1889 -----AlaAsnSerSerTrpSerHis----- 1895
DB 3253 ACCTCTGACTCTTATPACCAGGATACCTGTGGAAACATCACATTTACCTTTAAACAGGAG 3312
QY 1896 -----SerProValCysGluProValLysCysSerSerProGluAsnIleAsn 1913
DB 3313 ATGATGTACACAGTCTTACTACGGAGCACATTTGCAGTGAATTTGAGGAATTTGAAT--- 3369
QY 1914 GlyLysTyrIleLeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThr 1933
DB 3370 -----ATTGTGAAGATGTTCCGCTGAATATTCATCTACATGCTTGGAGGCT 3420
QY 1934 GlyTyrSerLeuGlnGlyProSerIleIleGluCysThrAlaSerGlyIleTrpAspArg 1953
DB 3421 TCCCTCTCAGCGAACAATGAAATACATGTGGCCATTTCTGCTGAAGATATACGGATGAT 3480
QY 1954 AlaProPro 1956
DB 3481 GGGAAACCG 3489

RESULT 19
US-09-855-722-6
; Sequence 6, Application US/09855722
; Patent No. 6638741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; NAME/KEY: sig_peptide
; LOCATION: (409)..(501)
; NAME/KEY: mat_peptide
; LOCATION: (502)..(4062)
US-09-855-722-6

Alignment Scores:
Pred. No.: 1,61e-52 Length: 4208
Score: 763.00 Matches: 288
Percent Similarity: 34.28% Conservative: 145
Best Local Similarity: 22.80% Mismatches: 400
Query Match: 3.82% Indels: 430
DB: 4 Gaps: 57

US-09-977-053-4 (1-3571) x US-09-855-722-6 (1-4208)

QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
DB 571 TCGCGCGCGCCCGGAACCCCGGAGACCGCAAGTGCACCCCGGACGAGTCTGCACATAC 630
QY 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
DB 631 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGCTCACGCGCGGGGGGCC 681
QY 869 GlyGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
DB 682 TGCAGCTTCGGCTCAGG----- 699
QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
DB 700 -----TCCACGCTGTCTATCGGGGGCAACACCTTCACCTCAAGGCCACCGCGGC 750
QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
DB 751 AACGACCGCAACCGCATCTGTCTTCCTTCAGTTTCGCTGGCGGAGTCTATACGTG 810
QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu----- 937
DB 811 CTGTGGAGCGGTGGGANTTCCAGTATATGACACCGTTCAACCTGACAGCTATATTGAAAG 870
QY 938 -----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
DB 871 GCTTCTACTCGGCGATGATCAACCCCGCGGAGTGGCAGACGCTGAGAGACACG 930
QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
DB 931 GCGGTGCCCCACTTTGAGTATCAGATCCGCGTGCACCTGTGATGACTACTATGGCTTT 990
QY 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
DB 991 -----GGCTGCAT 999
QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
DB 1000 AAGTTCGTGCGCCCGCAGAGATGACTTCTTTGGA----- 1032
QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
DB 1033 -----CACTATGCTGTGAC-----CAGAATGGCAACAAATTCG 1068
QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
DB 1069 ATGGAAGGTGGATGGGCCCGCAAGTGTAAACAGAGCTATTGTCCGACAGAGCTGCAGTCT 1128
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 1129 AAGCATGGGTCTTGCAAACTCCAGGTGACTCCAGGTGCTGAGGTGCGAG-----TAC 1173
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
DB 1174 GGCTGGCAGGCGCTG---TACTGTGATAAGTGCATCCACACCCCGGAGTCCGTCCACGGC 1230

1078 QY -----GlnProLysPheGlySerArgSerCysLeu 1087
1231 DB ATCTGTAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGGCGGCACGCTCTGT--- 1287
1088 QY SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
1288 DB -----GACAAGATCTCAATTACTGTGG 1311
1108 QY Val-----ProCysProGluGly-----LysPheSer 1116
1312 DB ACTCATCAGCGGTCTCAACGGGGAACTGTGACAAACACAGCCCTGACAAATATACG 1371
1117 QY ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
1372 DB TGTTC-----TGCCTGAGGGGTATTTCAGGACCCAACTCT--- 1407
1137 QY LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
1407 DB ----- 1407
1157 QY IleThrGluCysSerSerPheSerThrPheSerAlaAlaGluGluSerValPro 1176
1407 DB ----- 1407
1177 QY ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
1408 DB -----GAAATTGCTGAG-----CACGCC 1425
1197 QY CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
1426 DB TGCCTCTGATCCCTGTCTCAACAGGCGCAGCTGTAAAGAGACCTCCCTGGGGCTTTGAG 1485
1217 QY CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGlyCysSer 1236
1486 DB TGTGAGTGTTCCTCCAGCGCTGCAGCGGCCACATGCTCTCAACAACATTGATGACTGTCT 1545
1237 QY ProLeuProCysLeuAsnAnglyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
1546 DB CCTAATAACTGTTCCTCCAGCGGGCACCTGCCAGCCTGGTTAAACGGATTAAAGTGTGTG 1605
1257 QY CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSer 1276
1606 DB TGCCTCCAGTGGACTGGGAAACCTGCAGCTAGATGCATATGATGATGAGGCCAAA 1665
1277 QY ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
1666 DB CCTTGTGTAAACGCCAAATCCTGTAAAGAACTCTCATTCAGCTACTACTGCGACTGTCTT 1725
1297 QY LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
1726 DB CCGCGCTGGATGGTTCAGAAATTCAGCAATAATAATTAATGACTGCTTGGCCAG---TGT 1782
1317 QY LeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGly 1336
1783 DB CAGAAAGACCCCTCTCTCGGATTTGGTTAAAGTTATCGCTGTATCTGTCCACTGGC 1842
1337 QY PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
1843 DB TATGCAGCGCATCATCTGTGAGAGAGACATCGATGAATGTGCCAACCCTGTGTTGAAT 1902
1357 QY GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
1903 DB GGGGGTCACTGTTCAGATGAATAACACAGATTCCAGTGTCTGTGTCTCCACTGGTTTCTCT 1962
1377 QY GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
1963 DB GGAACCTCTGTAGCTGGACATCGATTATGTGAGCTTAATCTCTCCAGACGGTGGC 2022
1397 QY ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
2023 DB CAGTGTCTACAAACCGTGCAGTGACTATTCTTCAAGTGCCTCCAGGAGCTATGAGGGCAAG 2082

1417 QY ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
2083 DB AACTGCTCACACTGAAGAGACCACTGCCGCACGACCCCTGTGTGAAGTG----- 2130
1437 QY GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
2131 DB -----ATTGAC-----AGCTGCACAGTGGCC 2151
1457 QY MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
2152 DB ATGGCTTCCACAGC-----ACACCTGAAGGGTGCCTATATTTTCCTCC 2196
1475 QY AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
2197 DB AAC-----GCTGTGGTCTCTCAC 2214
1495 QY GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAla 1514
2215 DB GCGAAG-----TGCACAGATCAGTCCGGAGGCAAAATTC----- 2247
1515 QY IleThrTyrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
2248 DB ACCTGTGACTGTAAACAAGGCTTCACGGGAACATATCTGCCATGAATAATTAATGAC 2304
1535 QY GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGly 1554
2305 DB TGTGAGAGCAAC-----CCTGTAGAAACGGTGGCAGCTTGCATCATGCT 2349
1555 QY GlnGluAsnLysLysGlyGlyPheSerProAlaGluSerPheValGlySerIle 1574
2350 DB GTCAACTCTCTACAG-----TGCATCTGTAGTACGCTGGGAGGGCGCTAC 2397
1575 QY SerGlnLeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
2398 DB TGTAAACCAATATTAATGACTGCAGCCAGAACCCC----- 2433
1595 QY SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
2434 DB TGCACCAAT----- 2442
1615 QY IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
2443 DB -----GGGGGCGACGTGCGGACCTGGTCAATCACTTCTACTGTGACTGTAAAAATGGG 2496
1634 QY LeuGlyCysSerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
2497 DB TGGAAAGGAAGACCTCCCACTCAGCTGACAGTCACTGTGATGAGGCCACCTGCAACAC 2556
1654 QY ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
2557 DB GTTGGGACCTGTATGATGAGGGGATGCTTTTAAG----- 2592
1672 QY CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
2593 DB TGCATGTGCTCTGGCGCTGGGAAGGAACAACCTGTAACATAGCCGGAACAGTAGCTGC 2652
1692 QY ValProProLeuLeuAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThr 1711
2653 DB CTGCCCAACCCCTGCCATTAATGGG----- 2676
1712 QY ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
2677 DB -----GGCACATGTGTGTGTCACGGCGAGTCTTTACGTGCGTCC 2715
1732 QY AspAsnGlySerTyrAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
2716 DB TGCAGGAGAGCTGGAGGGG-----CCCATCTGTGCTCAGAAATACC----- 2757
1752 QY GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
2758 DB AATGACTGCACCCCTCATCTCCCTGTTTACAAACAGCGCACCTGTGTGTGAGAGAAC 2814
1767 QY SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779

Db 2815 TGGTACCGGTGCAATGTGCCCCGGGTTTGTCTGGGCCGACGTGCAGATAAACATCAAT 2874
Qy 1780 -----GlyLysAsnCysAlaGluProLe----- 1787
Db 2875 GAATGCCAGTCTTCACCTTGTGCTTGTGAGCGACCTGTGTGATGATGATGCTAC 2934
Qy 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluLeuThrVal 1807
Db 2935 CGGTGTGTGCTGCT-----CCAGGCGACAGTGTGCTCCAA----- 2970
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
Db 2971 -----TCCAGGAA-----GTTTCAGGAGA 2991
Qy 1828 ThrCysLeuGluSerGlyLysTrpAsnHisLeuIlePro----- 1840
Db 2992 CCTTGATCACCATTGGGAGT-----GTGATACAGATGGGCCCAATGGGATGAT 3042
Qy 1841 -----TyrCysLysAlaValSerCysGly 1848
Db 3043 GACTGTAAATACCTGCCAGTCCCTGAATGACCGATCGCTGCTCAAGGTCTGTGTGTC 3102
Qy 1849 ---LysProAlaIleProGluAsnGly-----CysIle 1858
Db 3103 CCTGCGACCTGCTGCTCCAAAGGGGACAGCGAGTGCCTCCAGCGCGGAGAGCTGCATC 3162
Qy 1859 -----GluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys 1875
Db 3163 CCCATCCTCGACGACCGAGTCTC-----GTCCACCCCTGC----- 3198
Qy 1876 GlyTyrThrLeuAlaGlyAspLysGluSerSerCysLeu----- 1888
Db 3199 -----ACTGGTGTGGCGAGTGTGCTGCTCTCCAGTCTCCAGCGCGTGAAGACAAAGTGC 3252
Qy 1889 -----AlaAsnSerSerTrpSerHis----- 1895
Db 3253 ACCTCTGACTCTTATACAGGATACCTGTGCGAATCATCATTTACCTTTACAGGAG 3312
Qy 1896 -----SerProProValCysGluProValLysCysSerSerProGluAsnIleAsn 1913
Db 3313 ATGATGTACACAGGTCTTACTACGAGGACATTTGCGAGTGAATGAGGAATTGAAT--- 3369
Qy 1914 GlyLysTyrIleLeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThr 1933
Db 3370 -----ATTGTGAAGAATGTTTCCGCTGAATATTCAATCTACATCGCTTGGAGCCT 3420
Qy 1934 GlyTyrSerLeuGlnGlyProSerIleIleGluCysThrAlaSerGlyIleTrpAspArg 1953
Db 3421 TCCCTTACGGAACAAATACATGATGGCCATTTCTGTGAGATATACGGGATGAT 3480
Qy 1954 AlaProPro 1956
Db 3481 GGGAAACCG 3489

RESULT 20

US-08-110-158-3
; Sequence 3, Application US/08110158
; Patent No. 5605821
; GENERAL INFORMATION:
; APPLICANT: McEver, Rodger P.
; TITLE OF INVENTION: Expression Control Sequences of the
; TITLE OF INVENTION: P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/110,158
FILING DATE: 19930820
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/320,408
FILING DATE: 08-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3142 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-110-158-3

Alignment Scores:
Pred. No.: 1,056-52 Length: 3142
Score: 762.50 Matches: 311
Percent Similarity: 32.49% Conservative: 140
Best Local Similarity: 22.41% Mismatches: 421
Query Match: 3.82% Indels: 517
DB: 1 Gaps: 60

US-09-977-053-4 (1-3571) x US-08-110-158-3 (1-3142)

Qy 1556 GluGlnAspLysLysGlyGlyPheSerProAlaGluSerPheValGlySerIleSer 1575
Db 20 CAGCAGATCATCAGAGAGATGGCCAACTGCCAATAGCCATCTTGTACAGATTCCTCA 79
Qy 1576 GlnLeuAsnLeuTrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSer 1595
Db 80 GAGAGTGGTCTTGGAAATTTCCAACTCTTGTCTCAGTGGCCCTGATCTCTGAATTAAC 139
Qy 1596 CysProGlu-GluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGly11 1615
Db 140 AAACCCAGAAAGAA-----GTGGCAGCATGAGCTTATCATTTACAGACAAA 184
Qy 1615 eValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArgLeuG1 1635
Db 185 AGCATACTCATGGAATATTTCCGTAAA-----TACTGCCAG----- 221
Qy 1635 yGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAs 1655
Db 222 -----AA 223
Qy 1655 nLeuPheCysAspProGlyPheGlnLeuValGly-----AsnProValGlnTy 1671
Db 224 TCGCTACACAGAC-----TTAGTGGCCATCCAGATAAAATGAAATTTGATTA 271
Qy 1671 rCys-----LeuAsnG1 1675
Db 272 CTCTAATAAGTCTTACCTCTACTACAGCTCTCTACTGTGGATGGGATCCGAAAGACAA 331
Qy 1675 nGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProProPr 1695
Db 332 TAAGACATGGACATGG-----GTGGACACCAAAAAGGC 364
Qy 1695 oLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrG1 1715
Db 365 TCTCACCAC-----GAGGCTGAGAAC---TGGGCTGATATGAACTTACACAA 412
Qy 1715 nCysAsnAsn-----GlyTyrTyrLeuLeuGlyAspSerArgMetPheCysTh 1731


```

Db      413  AGGAACACAGGAGCTGCTGGAGATATACATCAAGAGTCCGTCAGCC----- 461
Qy      1731  rAspAenGlySerTrpAenGly-----ValSe 1740
Db      462  -----CCTGGCAAGTGGAAATGATGAGCACTGCTTGAAGAAAAAGACGACGATTTGTTATCAC 517
Qy      1740  rProSerCysLeuAaspValAaspGluCysAlaValGlySerAaspCysSerGluHisAlaSe 1760
Db      518  AGCTCTCTCCAGGACATG-----TCTCGAGCAAAACAGGAGA 556
Qy      1760  rCysLeuAenValAaspGlySerTrpAenGlySerCysValProProTrpThrGlyAaspG1 1780
Db      557  GTGCTCGAGACCATCGGAGACTACACCTGCTCTGTTACCTCGGATTCATGGCCGAGA 616
Qy      1780  y-----LysAenCy 1783
Db      617  ATGTGAATACGTGAGAGAGTGTGGAGAACTTGAGCTCCCTCAACACAGTGCTCATGAACCTG 676
Qy      1783  sAlaGluProIleLysCysAlaProGlyAenProGluAenGlyHisSerSerGlyG1 1803
Db      677  CAGCCACCTCTG-----GGAAC 695
Qy      1803  uIleTyThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyThrGlnLeuMetG1 1823
Db      696  ----TTCTCTTTAACTCGCAGTGCAGCTTCCACTGCAGCTGACGGGTACCAAGTAAATGG 751
Qy      1823  yValThrLysIleThrCysLeuGluSerGlyGluTrpAenHisLeuIleProTyCysLy 1843
Db      752  GCCCAAGACTGGAATGCTTGCTTCTGGAATCTGGACAAATAAGCCTCCACAGTGT 811
Qy      1843  sAlaValSerCysGlyValProAlaIleProGluAenGlyCysIleGluLeuAlaPh 1863
Db      812  AGCTGCCAGTGCACCCCTGAGAGATTCTGACGAGGA----- 851
Qy      1863  eThrPheGlySerLysValThrTyArgCysAenLysGlyTyThrLeuAlaGlyAaspLy 1883
Db      851  ----- 851
Qy      1883  sGluSerSerCysLeuAlaAenSerSerTrpSerHisSerProProValCysGluProVa 1903
Db      852  -AACATGATCTGCCTT-----CATCT 872
Qy      1903  lLysCysSerSerProGluAenIleAenAenGlyLysTyIleLeuSerGlyLeuThrTy 1923
Db      873  -----GCAAAAGCATTCAGCA 889
Qy      1923  rLeuSerThrAlaSerTySerCysAaspThrGlyTySerLeuGlnGlyProSerIle1 1943
Db      890  TCAGTCTAGCTGACGCTTCAGTTGTGAAGAGGATTTGCATTAGTTGGACCGGAAGTGT 949
Qy      1943  eGluCysThrAlaSerGlyIleTrpAaspArgAlaProProAlaCysHisLeuValPheCy 1963
Db      950  GCAATGCACAGCTCGGGGGTATGGACAGCCCCCAGCTGTGTAAAGCTGTGCAGTG 1009
Qy      1963  sGly-----GluProProAla-----IleLysAaspAlaValIleThrGlyAenAenPh 1979
Db      1010  TCAGCACCTGGAGGCCCGCCAGTGAAGAACCATGGACTGTGTTCATCCGCTCACTCTT 1069
Qy      1979  eThrPheArgAenThrValThrTyThrCysLysGluGlyTyThrThrLeuAlaGlyLeuAe 1999
Db      1070  TGCCTATGCTCCAGCTGCATAATTTGAGTGCCAGCCCGCTACAGAGTGGGGCTTGA 1129
Qy      1999  pThrIleGluCysLeuAlaAaspGlyLysTrpSerArgSerAaspGlnCysLeuAlaVa 2019
Db      1130  CATGCTCCGCTGCATTTGACTCTGGACACTGCTGCACCCCTTGCCAAACCTGTGAGGCTAT 1189
Qy      2019  lSerCysAasp-----GluProProIleValAaspHisAlaSerProGluThrAlaHisAr 2037
Db      1190  TTCGTGTGAGCGCTGGAGAGTCTCTC-----CACGAGAGCATGGAT----- 1232
Qy      2037  gLeuPheGlyAaspIleAlaPheTyTyCysSerAaspGlyTyThrSerLeuAlaAaspAenSe 2057

```

```

Db      1233  -----TGCTCT----- 1238
Qy      2057  rGlnLeuLeuCysAenAlaGlnGlyLysTrpValProProGluGlyGlnAaspMetProAr 2077
Db      1238  ----- 1238
Qy      2077  gCysIleAlaHisPheCysGluLysProProSerValSerTySerIleLeuGluSerVa 2097
Db      1239  -----CCATCCTTG----- 1247
Qy      2097  lSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheVa 2117
Db      1248  -----AGAGCGCTTTCAGTATGACACCACTGTAGCTTCGCTGTCTGAAGGTTTCAT 1300
Qy      2117  lLeuAenThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAen-----ProSerPr 2136
Db      1301  GCTGAGAGGAGCCGATATAGTTGCGTGTGATACTTGGAGAGTGGACAGCAGCACCGCCCC 1360
Qy      2136  oMetSerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAenGlyTy 2156
Db      1361  AGTC-----TGTCAGCTTTGCAGTGCCAGGATCTCCA-----GTTCCAATAGAGC 1408
Qy      2156  rAlaSerGlySerAenTySerPheGlyAla-----MetValAlaTySe 2171
Db      1409  CCGGTGAACTGCTCCACACCCCTTCGTCCTTTAGGTACCAGTCAGTCTGCAGCTTCAC 1468
Qy      2171  rCysAenLysGlyPheTyIleLysGlyGluLysLysSerThrCysGluAlaTrpGlyG1 2191
Db      1469  CTCGAATGAAGGCTTCTCTCTGGTGGAGCAAGTGTCTACAGTCTGCTGCTACTGGAA 1528
Qy      2191  nTrpSerSerProIleProThrCysHisProValSerCysGlyGluProProLysValG1 2211
Db      1529  CTGGAATTCGTTCTCCAGAAATGCCAAGCCATTCCCTGCACACCTTTGCTAAGCCCTCA 1588
Qy      2211  uAenGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyGlnCy 2231
Db      1589  GAATGGA----- 1595
Qy      2231  sAenProGlyTyLysSerValGlySerProValPheValCysGlnAlaAenArgHisTr 2251
Db      1595  ----- 1595
Qy      2251  pHisSerGluSerProLeuMetCysValProLeuAaspCysGlyLysProProIleG1 2271
Db      1596  -----ACATGACCTGTGT-----CAACCTCTTG 1621
Qy      2271  nAenGlyPheMetLysGlyGluAenPheGluValGlySerLysValGlnPhePheCysAs 2291
Db      1622  AAGTTCAGCTTATAA-----TCCACATGTCAATTCATCTGTGA 1660
Qy      2291  nGluGlyTyThrGluLeuValGlyAaspSerSerTrpThrCysGlnLysSerGlyLysTrpAs 2311
Db      1661  CGAGGATATCTTTGCTGGACAGAAAGATTGGATTGTACTCGATCGGAGACCTGGAC 1720
Qy      2311  nLysLysSerAenProLysCysMetProAlaLysCysProGlu-----ProProLe 2328
Db      1721  A---GACTCCCCCACTGTGTGAAGCCATCAAGTGCCAGAACTCTTTGCCCCCAGAG-- 1775
Qy      2328  uLeuGluAenGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThr----- 2346
Db      1776  -----CAGGCGAGCTGTGATTGTTCTGACACTCTGTGGAGAAATTCAAATTTGGCTCCACCTG 1831
Qy      2347  -----PheSerCysLysGlyHisValLeuGlnGlyProSerValLeuLysCysLeuPr 2365
Db      1832  TCATTTCTCTTGTAAACAATGGCTTTAAGCTGGAGGGGCCCAATATGTGGATGCACAC 1891
Qy      2365  oSerGlnGlnTrpAenAaspSerPheProValCysLysIleValLeuCysThrProProPr 2385
Db      1892  TTCTGGAAGATGTGCTAGCTACTCCACCACTGCAAAAGGCATAGCATCACTTCTCTACTCC 1951
Qy      2385  oLeuIleSerPheGlyValProIleProSer----- 2395
Db      1952  A-----GGGTTCGAATGTCCAGCCCTCACCACTCTCTGGGCGAGGGAACCATGTA 1999

```

QY 2396 -----SerAlaLeuHisPheGlySerThrValLysTyrSerCysValG1 2410
DB 2000 CTGTAGGCATCATCGGAAACCTTTGGTTTAAATACCACTGTGTACTTTGGCTGCAAGC 2059
QY 2410 yGlyPheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSe 2430
DB 2060 TGGATTTCACACTATAGAGAGACGACCTCTCAGTCGACGACCTTCAGAGCAATGACAGC 2119
QY 2430 rProLeuProGluCysValProValGluCysProGlnProGluGluLeuProAsnGlyI1 2450
DB 2120 AGTAACTCCAGCATGCAGAGCTGTCAATGCTCAGAA----- 2156
QY 2450 eIleaspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPh 2470
DB 2157 -----CTACATGTTAATAAGCCA----- 2174
QY 2470 eGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyLy 2490
DB 2175 ----ATAGCATGAACCTGCTCAACCTC-----TGG----- 2201
QY 2490 sProThrCysLysAlaIleGluCysLeuLysProLysGluLeuLeuAsnGlyLysPheSe 2510
DB 2202 -----GGAACTTCAG 2212
QY 2510 rTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLe 2530
DB 2213 T-----TATGATCAATCTGCTCTTTTCCATGCTCTAGAGGGCCAGTTACT 2257
QY 2530 uGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValaspAlaProSe 2550
DB 2258 TAATGGCTCTGCAAAACAGCATGCTCAAGAGAAATGGCCACTGTCTCACTACCGTCCCAAC 2317
QY 2550 rCysAsnAlaIleHisCysAspSerProGlnProIleGlu-----AsnG1 2565
DB 2318 CTGCCAAGCAGG-ACCAATGACTATCCAGGAGCCCTGACTTACTTTGTTGGAGCGGTGG 2376
QY 2565 yPheValGluGlyAlaAspTyrSerTyrGlyAlaIlelleIleTyrSerCysPheProG1 2585
DB 2377 CTCTTACAATAGTCTGATAATGGTGGGACGCT-----CCTGG 2415
QY 2585 yPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerSer11 2605
DB 2416 CTTT-----GCTAAGAAGCGTTTCAGACA-----AAAGATGATGGAA----- 2455
QY 2605 eProThrCysMetProIleAspCysGlyLeuProProHis----- 2618
DB 2456 -----ATGCCCTTGTAATCTCTCAGCCACTAGAACATATGGAGTTTITACAA 2505
QY 2618 ----- 2618
DB 2506 AGCTGCATTTGACCCGAGTCTTAAGGTTTCCATAAAACCCATCAATCAAGACATGG 2565
QY 2619 -----Ileasp----- 2620
DB 2566 AATTACCTTAGATTAGTCTTGACACGAGCTGTGGACCGCTCTGGACCAACCCCTGTTTC 2625
QY 2621 ----PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAs 2639
DB 2626 CTGAGTTTGGGATTGTTGGTACAAATCTCAATCTCACT----- 2665
QY 2639 pMetMetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLy 2659
DB 2666 -----ACCACCCCTTCCTGTCCTC----- 2683
QY 2659 sThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyTh 2679
DB 2684 -----ACCTCTCTCTCTTCTGTAACACAGCCACAGGACGAGGAGCAATG 2730
QY 2679 rMetValSerTyrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCy 2699
DB 2731 TTTCTGCAGTAGTCTCTGTTGCTTACTCCTGTTACTTGAATACCACTG-----AA 2784

QY 2699 sGlnGluAspGlyThrTrpAsn-----GlySerAlaProSerCysIleSerileG1 2716
DB 2705 CCAAGAGAC-----TGAGCATCTGACTCACAGAGACCACTGTGGAGAAATAAA 2838
QY 2716 uCysAspLeuProThrAlaProGlu-AsnGlyPheLeuArgPheThrGluThrSerMetG 2736
DB 2839 AATACCTCTTATTATTTTGAATTGAAGGAAGTTTCTCCACTTTGTTGGAAGCAGGTGG 2898
QY 2736 lYserAlaVal---GlnTyrSerCysLys-ProGlyHisIleLeuAlaGlySerAspLeu 2754
DB 2899 CATCTCTAATTGAAGAAATTCCTCTAGCATCTCTGAGATCTCCAGTGGT----- 2949
QY 2755 ArgLeuCys-LeuGluAsnArgLysTrpSerGlyAla-----SerProAr 2769
DB 2950 -----TGCTGTGATGAGGCTCTTGACCTCTCTGAGGCTTCCAGAGATCTCT 3003
QY 2769 gCysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerileysGlySe 2789
DB 3004 GGATGGCACCAGAGCTGCAAGAGCCCAAGATCAAGCTAGAAGGCCACATG-----TC 3057
QY 2789 rAsnTyrThrTyrLeu 2794
DB 3058 ACCGTGGACCTTCCTG 3073

RESULT 21

US-09-023-655-1090
; Sequence 1090, Application US/09023655
; Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,655
FILING DATE: HERewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeiler, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1090:

SEQUENCE CHARACTERISTICS:

LENGTH: 3142 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g183390

US-09-023-655-1090

Alignment Scores:

Pred. No.:	1.05e-52	Length:	3142
Assignment Score:	Score:	Matches:	311
	762.50	Conservative:	140
	Percent Similarity:	Mismatches:	421
	Best Local Similarity:	Indels:	517
	Query Match:	Gaps:	60
			4
			DB:

US-09-977-053-4 (1-3571) x US-09-023-655-1090 (1-3142)

1556	GlulnAspValysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSer	1575
20	CAGCAGAGTCCACGAGGAGATGGCCAACTGCCAAATAGCCATCTTGTACCAGAGATTCCA	79
1576	GlnLeuAsnLeuThrAspTyrValLeuSerProGlnGlnVallysSerLeuAlaThrSer	1595
80	GAGAGTGGTCTTTGGAAATTTCCCAACTCCTTTTGCTTCAGTGCCCTGATCTCTGAACTAAC	139
1596	CysProGlu-GlnLeuSerIysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyI1	1615
140	AAACCAGAGAA-----GTGGCAGCATGGACTTATCATTCACGACACAA	184
1615	evalGlyIysVallysIleAspSerIysSerIlePheCysSerAspCysProArgLeuG1	1635
185	AGCATACTCATGGAATATTTCCCGTAA-----TACTGCCAG-----	221
1635	yGlySerValProHisLeuArgThrAlaSerGluAspLeuIysProGlySerIysValas	1655
222	-----	223
1655	nLeuPheCysAspProGlyPheGlnLeuValGly-----AsnProValGlnTy	1671
224	TCGCTACACAGAC-----TTAGTGGCCATCCAGATATAAATAATGAAATTGATTA	271
1671	rCys-----LeuAsnG1	1675
272	CCTCAATAAGTCTCCTACTACTACAGCTCTACTACTGGATTGGGATCCGAAAGACAA	331
1675	nGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGlyValProPro	1695
332	TAAACATGTGCACATGG-----GTGGGAACCAAAAGGC	364
1695	oLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrG1	1715
365	TCTCACCAC-----GAGGCTGAGAAC---TGGGCTGATAATCACTTAACACAA	412
1715	nCysAsnAsn-----GlyTyrTyrLeuLeuGlyAspSerArgMetPheCysTh	1731
413	AAGGAACAACGAGGACTCGCTGGAGATATATACAGAGTCCGTCAGCC-----	461
1731	rAspAsnGlySerTrpAsnGly-----Valse	1740
462	-----CCTGGCAAGTGGAAATGATGAGCACTGCTTGAGAAAAGCACGCATTTGTGTACAC	517
1740	rProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlase	1760
518	AGCCTCTCTCCAGACATG-----TCTGTCAGACAAACAAGGAGA	556
1760	rCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspG1	1780
557	GTGCCTCGAGACCATCGGAACTACACCTGCTCTCTGTATCCCTGGATTCTATGGGCCAGA	616
1780	y-----LysAsnCy	1783
617	ATGTGAATACGTGAGAGATGTGAGAACTTTGAGCTCCCTCAACACGTGCTCATGAAC	676
1783	salAGluProIleIysCysIysAlaProGlyAsnProGluAsnGlyHisSerSerGlyG1	1803
677	CAGCCACCCTCTG-----GGAAAC-----	695
1803	uIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetG1	1823
696	-----TTCTCTTTTAACTCGCAGTGCACCTTCCATCTGCATCTCAAGGGTACCAAGTAATGG	751

Qy	1823	yValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyrCysLy	184
Db	752	GCACAGACAGCTGGAAATGCTTGGCTTCTGGAAATCTGGACAATTAAGCCTCCACAGTGT	811
Qy	1843	eAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGluLeuAlaPh	1863
Db	812	AGCTGCCAGTGCACCCCTGAAGATTCTCGAACGAGGA	851
Qy	1863	eThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLy	1883
Db	851	-----	851
Qy	1883	sGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCysGluProVa	1903
Db	852	-AACATGATCTGCCTT-----CATCTC-----	872
Qy	1903	lLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeuThrTy	1923
Db	873	-----GCAAAAGCATTTCCAGCA	883
Qy	1923	rLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleI	1943
Db	890	TCAGTCTAGCTGACGCTTTCAGTTGTGAAGAGGATTTGCATTAGTTAGTCCGACCGGAAGTGGT	949
Qy	1943	eGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCy	1963
Db	950	GCATGCACAGCCTCGGGGTATGGACAGCCCGCCAGCTGTGTAAAGCTGTGCAGTG	1009
Qy	1963	sGly-----GluProProAla-----IleLysAspAlaValIleThrGlyAsnAsnPh	1979
Db	1010	TCAGCACCCTGGAGCCGCCAGTGAAGAACCATGACTGTGTTCATCCGCTCACTGCTTT	1069
Qy	1979	eThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAs	1999
Db	1070	TGCCTATGGCTTCAGCTGCCAAATTTAGTGGCCAGCCCGCTACAGATGAGGGCTTGGGA	1129
Qy	1999	pThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaVa	2019
Db	1130	CATGCTCCGCTGCATTGACTCTGCACACTGGTCTGTCACCCCTTGCCAACTGTGAGGCTAT	1189
Qy	2019	lSerCysAsp-----GluProProIleValAspHisAlaSerProGluThrAlaHisAr	2037
Db	1190	TTGCTGTGAGCCGCTGGAGAGTCTCTGTC-----CACGGAAGCATGGAT-----	1232
Qy	2037	gLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSe	2057
Db	1233	-----TGCTCT-----	1238
Qy	2057	rgInLeuLeuCysAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProAr	2077
Db	1238	-----	1238
Qy	2077	gCysIleAlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerVa	2097
Db	1239	-----CATCTCTG-----	1247
Qy	2097	lSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheVa	2117
Db	1248	-----AGAGCGTTTCAGTATGACACCACTGTAGCTTCGGCTGTGCTGAAGGTTTCAT	1300
Qy	2117	lLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsn---ProSerPr	2136
Db	1301	GCTGAGGAGGACCGATATAGTTCCGTTGTGATACTTGGGACAGTGGACGACACAGCCCC	1360
Qy	2136	oMetSerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTy	2156
Db	1361	AGTC-----TGTCAAGCTTGGAGTCCAGATCTCCCA---GTTCCAAATGAGGC	1408
Qy	2156	rAlaSerGlySerAsnTyrSerPheGlyAla-----MetValAlaTyrSe	2171
Db	1409	CCGGTGAACACTGCTCCACCCCTTCGGTGCTTTTAGTACAGTCAGTCTGCGAGCTTCAC	1468

Qy	2171	rCysAsnLysGlyPheTyrIleLysGlyGluLysLysSerThrCysGluAlaThrGlyG	2191
Db	1469	CTGCAATGAAGGCTTGCCTCGGTGGGAGCAGTGTGTACAGTCTGTGGCTACTGGAAA	1528
Qy	2191	nTrpSerSerProIleProThrCysHisProValSerCysGlyGluProProLysValG	2211
Db	1529	CTGGAATTCTTCTCCACGAATGCCAAGCCATTCCCTGCACACCTTGTCAAGCCCTCA	1588
Qy	2211	uAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCy	2231
Db	1589	GAATGGA-----	1595
Qy	2231	sAsnProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTr	2251
Db	1595	-----	1595
Qy	2251	pHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIleG	2271
Db	1596	---ACAAATGACCTGTGTT-----CAACCTCTTGG	1621
Qy	2271	nAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAs	2291
Db	1622	AAAGTCCAGTTATAA-----TCCACATGTCATTCACTGTGCA	1660
Qy	2291	nGluGlyTyrGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAs	2311
Db	1661	CGAGGATATCTTTGTCTGGACCAAGAAAGTTGATTGTACTCGATCGGAGCGTGGAC	1720
Qy	2311	nLysLysSerAsnProLysCysMetProAlaLysCysProGlu-----ProProLe	2328
Db	1721	A---GACTCCCCACCAATGTGTGAAGCCATCAAGTGCCTGAACTCTTGTCCCAAG--	1775
Qy	2328	uLeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThr----	2346
Db	1776	---CAGGCGACCTGGATTGTTCTGCACACTCGTGGAGATTCAATGTTGGCTCCACCTG	1831
Qy	2347	---PheSerCysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuPr	2365
Db	1832	TCATTTCTTTGTAACAATGCTTTAAGCTGGAGGGGCCCAATAATGTGGAATGCACAAC	1891
Qy	2365	oSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCysThrProProPr	2385
Db	1892	TTCTGGAGAGATGTCAGTACTCACCACCACTGCAAAAGGCATAGCATCTTCTACTCC	1951
Qy	2385	oLeuIleSerPheGlyValProIleProSer-----	2395
Db	1952	A-----GGTGTCAATGTCAGGCCTCACCTCCCTGGGAGGGAACCATGTA	1999
Qy	2396	-----SerAlaLeuHisPheGlySerThrValLysTyrSerCysValG	2410
Db	2000	CTGTAGGCATCATCGGGAACCTTTGGTTTTAATACCACCTGTGTACTTGGCTGCAACGC	2059
Qy	2410	yGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSe	2430
Db	2060	TGGATTACACTCATAGGAGACAGCACTCTCAGCTGCAGACCTTCAGGACAATGGACAGC	2119
Qy	2430	rProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIl	2450
Db	2120	AGTAACTCCAGTCAGAGCTGTGAATGCTCAGA-----	2156
Qy	2450	eIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPh	2470
Db	2157	-----CTACATGTTAATAAGCCA-----	2174
Qy	2470	eGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLy	2490
Db	2175	---ATACCGATGAAGTCTCTCAACCTC-----TGG-----	2201
Qy	2490	sProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSe	2510
Db	2202	-----GGAAACTTCAG	2212
Qy	2510	rTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLe	2530

RESULT 22
US-09-199-865-2

```
; Sequence 2, Application US/09199865
; Patent No. 6433138
; GENERAL INFORMATION:
; APPLICANT: Zimin, Ann B.
; APPLICANT: MacIag, Thomas
; APPLICANT: Wong, Michael K.K.
; APPLICANT: Pepper, Michael S.
; APPLICANT: Montesano, Roberto
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS
; FILE REFERENCE: 0036-1U1
; CURRENT APPLICATION NUMBER: US/09/199,865
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: 60/018,841
; EARLIER FILING DATE: 1996-05-31
; EARLIER APPLICATION NUMBER: PCT/US97/09407
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5458
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-199-865-2

Alignment Scores:
Pred. No.: 5,48e-52 Length: 5458
Score: 759.00 Matches: 288
Percent Similarity: 34.13% Conservative: 143
Best Local Similarity: 22.80% Mismatches: 402
Query Match: 3.80% Indels: 430
DB: 4 Gaps: 57

US-09-977-053-4 (1-3571) x US-09-199-865-2 (1-5458)
QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
DB 133 TGGCGGGCGCGCGGAGACCGCGGAGACCGGAGTGCACCGCGAGAGTGTGACACATAC 192
QY 851 ----LysTyrCysLeuGluThrAsnTyrAspTyrGluAsnGlyPhe-----AlaIleGlyPro 868
DB 193 TTCAAAAGTGTGCTCAAG-----GAGTATCAGTCCCGCGTCACCGCGCGGGGCGCC 243
QY 869 GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
DB 244 TGCAGCTTCGGCTCAGGG----- 261
QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
DB 262 -----TCCAGCGCTGTATCGGGGCAACACCTTCAACCTCAAGCGCAGCGCGGC 312
QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
DB 313 AACGACCGCAACCGCATCGTGTGCTTTCAGTTTCGCTCGCGGAGGCTCTATACGCTTG 372
QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluThrGlu----- 937
DB 373 CTGTGAGCGGTGGGATTCAGTAATGACACCGTTCACCTGACGATATTATTGAAAG 432
QY 938 -----AspGlnGlnArgLeuLeuGluGlnThrLeuGluThrIleThr 950
DB 433 GCTTCTCACTCGGGCATGATCAACCCCGCGGAGTGCAGCGCTGAGCAGACAGCAACAG 492
QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
DB 493 GCGGTTGCCCATTTGAGTATCATCGTCCGCGTACCTGTGATGACTACTACTATGCGCTTT 552
QY 966 GlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
DB 553 -----GCTGTAAAT 561
QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
DB |||||
```

```
DB 562 AAGTTCTGCGCGCCCGCAGAGATGACTTCTTTGGA----- 594
QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
DB 595 |||||-----CACTATGCTGTGTGAC-----CAGAAATGGCAACAAACTTGC 630
QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
DB 631 ATGGAAGCTGGATGGCGCCCGATGTAACAGAGCTATTTCGACAGAGGCTGCACTCT 590
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 691 AAGCATGGGTCTTGCACAACTCCAGGTGACTGTCAGGTGCCAG-----TAC 735
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
DB 736 GCGTGGCAAGGCGCTG---TACTGTGATAAGTGCATCCACACCGCGGAGTGGTCCACGCG 792
QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
DB 793 ATCTGTAATGAGCCCTGCGAGTGCCTCTGTGAGACCAACTGGGCGCGCAGCTCTGT--- 849
QY 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
DB 850 -----GACAAAGATCTCAATTACTGTGG 873
QY 1108 Val-----ProCysProGluGly-----LysPheSer 1116
DB 874 ACTCATCAGCGGTGTCTCAACGGGGGAACTTGTAGCAACACAGCGCCCTGACAAATATCAG 933
QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
DB 934 TGTTCC-----TGCCCTGAGGGGTATTTCAGGACCCCACTGT--- 969
QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
DB 969 ----- 969
QY 1157 IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValPro 1176
DB 969 ----- 969
QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
DB 970 -----GAAATTGCTGAG-----CACGCC 987
QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
DB 988 TGCCTCTCTGATCCCTGTCTCAACAGAGCGAGCTGTAAAGGAGACCTCTCGGCTTTGAG 1047
QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
DB 1048 TGTGAGTGTTCCTCCAGGCTGACCGCGCCCATGCTCTACAAACATTGATGACTGTCT 1107
QY 1237 ProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGlu 1256
DB 1108 CCTAATAACTGTTCCACGCGGGGACCTGCCAGGACCTGGTTAAGGATTTAAGTGTGTG 1167
QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer 1276
DB 1168 TGCCTCCACAGTGGAGTGGGAAACCGTCCAGTGTAGTGCATGCAATGAATGTGAGGCCAAA 1227
QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
DB 1228 CTTGTGTAAACGCGCAAAATCCTGTAAGATCTCAATGCCAGCTACTACTGCGACTGTCTT 1287
QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
DB 1288 CCGCGCTGGATGGTTCAGAAATGTGACATAAATATTAAATGACTGCTTGTGGCCAG---TGT 1344
QY 1317 LeuAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProGly 1336
DB 1345 CAGAAATGACGCGCTCTGTGCGGATTTGTTAAATGTTATCGCTGTATCGCTGTCTGCTGCGC 1404
```

```
QY 1337 PheLeuGlyThrArgCysGlyIysAsnValAspGluCysLeuSerGlnProCysIysAsn 1356
DB 1405 TATGACGGCGATCACTGTGAGAGAGACATCGATGAATGTGCAGCAACCCCTGTTTGAAT 1464
QY 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
DB 1465 GGGGTCACTGTCAAGATGAATCAACAGATTCAGTGTCTGTGTCCCACTGGTTCTCT 1524
QY 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
DB 1525 GGAACCTCTGTCACTGGACATGATATTGTGAGCCTAATCTCTGCAGAACGGTGCC 1584
QY 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
DB 1585 CAGTGTCAACACCGTCCAGTCACTATTCTGCAAGTGCCTCCGAGGAGCTATGAGGGCAAG 1644
QY 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
DB 1645 AACTGCTCACCTGGAAGACCACTGCCGCGACAGCCCTGTGAAGTG----- 1692
QY 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
DB 1693 -----ATTGAC-----AGCTGCACAGTGGCC 1713
QY 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
DB 1714 ATGGCTTCCACGAC-----ACACCTGGAAGGGGTGGCGGTATATTTCCTCC 1758
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
DB 1759 AAC-----GTCTGTGCTCTCTAC 1776
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTrpHisIleAla 1514
DB 1777 GGGAG-----TGCAAGACTCAGTCGGGAGGCAATTC----- 1809
QY 1515 IleThrTrpThrSerAlaAsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAsp 1534
DB 1810 ---ACCTGTGACTGTAAACAAGGCTTCACGGGAACATATCTGCCATGAAATAATTATGAC 1866
QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyGlyAlaLeuLeuGly 1554
DB 1867 TGTGAGAGCAAC-----CCTGTAGAACGGTGGCACTTGCATCGATGGT 1911
QY 1555 GlnGluGlnAspLysGlyGluGlyPheSerProAlaGluSerPheValGlySerIle 1574
DB 1912 GTCAACTCTACAG-----TGCACTGTGTAGTGACGGCTGGGAGGGGGCTAC 1959
QY 1575 SerGlnLeuAsnLeuTrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
DB 1960 TGTGAACCAATATTAAATGACTGCAGCCAGAACCC----- 1995
QY 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGly 1614
DB 1996 ---TGCCACAAAT----- 2004
QY 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
DB 2005 -----GGGGGCACGTGTCGGACCTGGTCAATGACTTCTACTGTGACTGTATAAATGGG 2058
QY 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
DB 2059 TGGAAAGGAACACCTGCCACTCACTGACGTGACGTGATGAGGCCACGTGCACACAC 2118
QY 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
DB 2119 GTTGCACCTGTATGATGAGGGGATGCTTTTAA----- 2154
QY 1672 CysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
DB 2155 TGCATGTGTGCTGGCGCTGGGAAGGAACAACCTGTGAACATAGCCCGGAACAGTAGCTGC 2214
```

RESULT 23

```
QY 1692 ValProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGlySerThr 1711
DB 2215 CTGCCAACCCCTGCCATATGGG----- 2238
QY 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
DB 2239 -----GGCACATGTGTGGTCAACAGCGGAGTCTTTTACGTGCGTC 2277
QY 1732 AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
DB 2278 TGCAGGAGGAGCTGGAGGGG-----CCCATCTGTGTCTCAGAATACC----- 2319
QY 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
DB 2320 ---AATGACTGCAGCCCTCATCCCTGTACACAGCGGCACTGTGTGATGGAGACAAC 2376
QY 1767 SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779
DB 2377 TGGTACCGGTGCAATGTGCCCGGGTTTGTGTGGCCCGGACTGCAGAAATAAACATCAAT 2436
QY 1780 -----GlyLysAsnCysAlaGluProIle----- 1787
DB 2437 GAATGCCAGTCTTCACTTGTGCTTGTGGAGCGACCTGTGTGGATGAGATCAATGGCTAC 2496
QY 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
DB 2497 CGGTGTGTCTGCCCT-----CCAGGCGACAGTGTGTGCCAAG----- 2532
QY 1808 GlyAlaAlaValThrPheSerCysGlnGlyTyrGlnLeuMetGlyValThrLysIle 1827
DB 2533 -----TGCCAGGAA-----GTTTCAGGGAGA 2553
QY 1828 ThrCysLeuGluSerGlyLutTrpAsnHisLeuIlePro----- 1840
DB 2554 CCTTGATCACCATCGGGAGT-----GTGATACCAGATGGGCCCAATGGGATGAT 2604
QY 1841 -----TyrCysLysAlaValSerCysGly 1848
DB 2605 GACTGTATATACCTGCCAGTCCCTGAATGACGGATCGCTCTCAAAAGTCTGTGTGTGGC 2664
QY 1849 ---LysProAlaIleProGluAsnGly-----CysIle 1858
DB 2665 CCTGACACTGTGCTGTCCACAAAGGGCACAGCGAGTGGCCCGGCGCAGAGTGCATC 2724
QY 1859 -----GluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys 1875
DB 2725 CCCATCTCTGGACGACCGAGTGTTC-----GTCCACCCCTGC----- 2760
QY 1876 GlyTyrThrLeuAlaGlyAspLysGluSerSerCysLeu----- 1888
DB 2761 ---ACTGTGTGGCGAGTGTGGTCTTCCAGTCTCCAGCGGTGAAGACAAAGTGC 2814
QY 1889 -----AlaAsnSerTrpSerHis----- 1895
DB 2815 ACCTCTGACTCTATTACCAGGATAACTGTGCGAACATCACATTTTACCTTTAAACAGGAG 2874
QY 1896 -----SerProValCysGluProValLysCysSerSerProGluAsnIleAsnAsn 1913
DB 2875 ATGATGTCCACAGGTCTTACTACGGAGCACATTTGCAGTGAATTTGAGGAATTTGAAT--- 2931
QY 1914 GlyLysTyrIleLeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThr 1933
DB 2932 -----ATTTGAGGAATGTTCCGCTGATATTCATCTACATCGCTCGGAGCCT 2982
QY 1934 GlyTyrSerLeuGlnGlyProSerIleIleGluCysThrAlaSerGlyIleTrpAspArg 1953
DB 2983 TCCCTTTCAGCCCAACAATGAATATCATGTGGCCATTTCTGCTGAAGATATACGGGATGAT 3042
QY 1954 AlaProPro 1956
DB 3043 GGGAAACCCG 3051
```

US-08-400-159-5
 ; Sequence 5, Application US/08400159
 ; Patent No. 5869282
 ; GENERAL INFORMATION:
 ; APPLICANT: Ieh-Horowicz, David
 ; APPLICANT: Henrique, Domingos M.P.
 ; APPLICANT: Lewis, Julian H.
 ; APPLICANT: Myat, Anna M.
 ; APPLICANT: Fleming, Robert J.
 ; APPLICANT: Artavanis-Tsakonas, Spyridon
 ; APPLICANT: Mann, Robert S.
 ; APPLICANT: Gray, Grace B.
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
 ; TITLE OF INVENTION: SERATE GENE AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/400,159
 ; FILING DATE: 07-MAR-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-029
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 6464 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 371..4027
 ; US-08-400-159-5

Alignment Scores:
 Pred. No.: 7,41e-52 Length: 6464
 Score: 759.00 Matches: 288
 Percent Similarity: 34.13% Conservative: 143
 Best Local Similarity: 22.80% Mismatches: 402
 Query Match: 3.80% Indels: 430
 DB: 2 Gaps: 57

US-09-977-053-4 (1-3571) x US-08-400-159-5 (1-6464)
 QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluGluAsnLeuThrLys 850
 DB 533 TCGCGCGCGCCCGGAGACCCCGGAGACCGGAGTGCACCCCGCAGGTGTGACATAC 592
 QY 851 ---LyTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
 DB 593 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGCGTCACGCGCGGGGGGCC 643
 QY 869 GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
 DB 644 TGCAGCTTGGCTCAGCG----- 661

QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
 DB 662 -----TCCAGCGCTGTATCGGGGGCAACACCTTCAACCTCAAGGCAGCGCGGC 712
 QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
 DB 713 AACGACCCGAACCGCATCGTCTTCAGTTTCCTGCGCGAGGTCTCTATACGCTTG 772
 QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluThrGlu----- 937
 DB 773 CTTGTGAGGCGGTGGATTCACAGTAATGACACCGTTCACCTGACATATTATTGAAAG 832
 QY 938 -----AsnGlnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
 DB 833 GCTTCTCCTCGGCGATGATCAACCCCGCAGTGGCAGCGCTGAAGCAGAACACG 892
 QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
 DB 893 GCGGTTGCCCATCTTGTGATATCAGATCCGCGTGACCTGTGATGACTACTACTATGGCTTT 952
 QY 966 GlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAspSer 985
 DB 953 -----GGCTGTAAT 961
 QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
 DB 962 AAGTTCTGCGCGCCCGCAGAGATGACTTCTTTGGA----- 994
 QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
 DB 995 -----CACTATGCTCTGTGAC-----CAGAATGCAACAAACTTGC 1030
 QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
 DB 1031 ATGGAAGCGCTGGATGGCGCCCGAATGTAAACAGAGCTATTTGCCGACAGGCTGCAGTCT 1090
 QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
 DB 1091 AAGCATGGGTCTTGCAAACTCCAGGTGACTGAGGTGCCAG-----TAC 1135
 QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
 DB 1136 GCGTGGCAAGGCGCTG---TACTGTGATAGTGTATCCACACCGCGGATGCGTCCACGCG 1192
 QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
 DB 1193 ATCTGTAATGAGCCCTGGCAGTGCCTCTGTGAGACCACTGGGCGCGCCTCTGT--- 1249
 QY 1088 SerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAlaCysGly 1107
 DB 1250 -----GACAAAGATCTCAATTACTGTGG 1273
 QY 1108 Val-----ProCysProGluGly-----LysPheSer 1116
 DB 1274 ACTCATCAGCGGTGTCTCAACGCGGGGAACTTTAGCAACACAGCCCTGCAAAATATCAG 1333
 QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
 DB 1334 TGTTC-----TGCCCTGAGGGGTATTCAGGACCACTGT--- 1369
 QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
 DB 1369 ----- 1369
 QY 1157 IleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluSerValPro 1176
 DB 1369 ----- 1369
 QY 1177 ProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGlu 1196
 DB 1370 -----GAAATTGCTGAG-----CACGCC 1387
 QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216

1308 TGCCTCTGATCCCTGTGTCACACAGGCGAGCTGTAAAGAGACTCCCTGGGCTTTGAG 1447
 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
 1448 TGTGAGTGTCCCGAGGCTGGACCGGCCCCACATGCTCTACAAACATTGATGACTGTCT 1507
 1237 ProLeuProCysLeuAsnAsnGlyValCysLeuAspLeuValGlyGluPheIleCysGlu 1256
 1508 CCTAATACTGTTCCACGGGGGACCTCCAGGACTCGTTAAGGATTTAAGTGTGTG 1567
 1257 CysProSerGlyTyrThrGlyGlnArgCysGluAsnIleAsnGluCysSerSer 1276
 1568 TSCCCCCACAGTGGAGCTGGGAACGTCAGGTTAGTGAATGCAATGCAATGTGAGGCCAAA 1627
 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
 1628 CCTGTGTAAACGCCAATCCTGTGAAGATCAATTGCCAGCTACTACTGCCAGCTGTCTT 1687
 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
 1688 CCCGCTGGATGGGTGAGAAATGTGACATAAATATTAAATGACTGCCCTGGCCAG---TGT 1744
 1317 LeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGly 1336
 1745 CAGAAATGAGCCTCTCTGCGGATTTGGTTATGGTTATGGCTGTATCTGTCTCCACTGGC 1804
 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
 1805 TATGAGGCGCATCACTGTGAGAGACATCGATGAATGTGCCAGCAACCCCTGTTTGAAT 1864
 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
 1865 GGGGTCACTGTGAGAAATCAACAGATTCAGTGTCTGTGTGCCACTGTTTCTCT 1924
 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
 1925 GGAACCTCTGTGAGTGGACATGATATTATGTGAGCCTAATCCTCGCCAGAAAGTGGCC 1984
 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
 1985 CAGTGTCAACACCGTCCAGTCACTATTCTGCAAGTGCCTGAGGACTATGAGGGCAAG 2044
 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
 2045 AACTGCTCACCTGAAGACCACTGCCCGCCAGACCCCTGTGAAGTG----- 2092
 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrp 1456
 2093 -----ATTGAC-----AGCTGCACAGTGGCC 2113
 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
 2114 ATGGCTCTCAACGAC-----ACACCTGAAGGGGTGGCGGTATATTCTCTCC 2158
 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTrpValLeuTyrValAsn 1494
 2159 AAC-----GTCTGTGTCTCTAC 2176
 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTrpHisIleAla 1514
 2177 GGAAG-----TGCAAGAGTCACTCGGAGGCAAAATTC----- 2209
 1515 IleThrTrpThrSerAlaAsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAsp 1534
 2210 ---ACCTGTGACTGTAAACAAGGCTTCAACGGGAACATATGTCATGAAATATTATGAC 2266
 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyAlaLeuValLeuGly 1554
 2267 TGTGAGAGGCAC-----CCTGTAGAAACGGTGGCACTTGCATCGATGT 2311
 1555 GlnGluGlnAspLysGlyGlyGlyPheSerProAlaGluSerPheValGlySerIle 1574

2312 GTCAACTCTCTACAG-----TGCATCTGTAGTACCGCTGGGAGGGGGCCTAC 2359
 1575 SerGlnLeuAsnLeuTrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
 2360 TGTGAACCAATATTAAATGACTGCAGCCAGAACCCC----- 2395
 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGly 1614
 2396 ---TGCCACAAT----- 2404
 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
 2405 -----GGGGGCGAGCTGCGGACCTGTGTCATGACTTCTACTGTGACTCTGTAATAATGGG 2458
 1634 LeuGlyGlySerValProHisIleuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
 2459 TGAAGAAGGAAGACCTGCCACTCACCTGACAGTCAGTGTGATGAGCCACGTGCACACAC 2518
 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
 2519 GGTGGCACCTGCTATGATGAGGGGATGCTTTTAAG----- 2554
 1672 CysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
 2555 TGCATGTGCTCTGGCGCTGGGAAGGACACACCTGTAAACATAGCCCGAAACAGTAGCTGC 2614
 1692 ValProProLeuGluAsnGlyPheHisSerAlaAspPheTyrAlaGlySerThr 1711
 2615 CTGCCAACCCCTGCCATAATGG----- 2638
 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
 2639 -----GGCAGATGTGTGGTCAACGGCGAGTCTTTACGTGGTGC 2677
 1732 AspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
 2678 TGCAAGGAAGGCTGGGAGGG-----CCCATCTGTCTCAGAAATACC----- 2719
 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
 2720 ---AATGACTGAGCCCTCATCTCTTACACGGCGGACCTGTGTGGATGGAGACAAC 2776
 1767 SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779
 2777 TGTACCGTGGAAATGTCGCCGGGTGTTGTCGGCCGAGCTGCAGNATAAATCAATCAAT 2836
 1780 -----GlyLysAsnCysAlaGluProIle----- 1787
 2837 GAATGCCAGTCTTCACTTGTGCTTTGGAGCGACCTGTGTGATGAGATCAATGGCTAC 2896
 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleTyrThrVal 1807
 2897 CGGTGTGTGCTGCCCT-----CCAGGCGACAGTGTGTGCCAAG----- 2932
 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
 2933 -----TGCCAGGAA-----GTTTCAGGGAGA 2953
 1828 ThrCysLysLeuGluSerGlyGluTrpAsnHisLeuIlePro----- 1840
 2954 CTTGATCATCATTGGGAGT-----GTGATACAGATGGGCGCCAAATGGATGAT 3004
 1841 -----TyrCysLysAlaValSerCysGly 1848
 3005 GACTGTATATCTGCCAGTCTGATGATGACCGATCGCTCTCAAGGTCTGTGTGGC 3064
 1849 ---LysProAlaIleProGluAsnGly-----CysIle 1858
 3065 CCTGACCTTGGCTCTCCACAAAGGCGACAGAGTGGCCCGAGCGGAGAGCTGCATC 3124
 1859 -----GluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys 1875
 3125 CCCATCTCTGGACGACGAGTGTCTTC-----GTCCACCCCTGC----- 3160

QY 1876 GlyTyrThrLeuAlaGlyAspLysGluSerSerCysLeu----- 1888
DB 3161 -----ACTGGTGTGGCGAGTGTCCAGTCTCCAGCGGTGAAGACAAAGATGC 3214
QY 1889 -----AlaAsnSerTirpSerHis----- 1895
DB 3215 ACCTCTGACCTCTATTACACAGGATACTGTGGACATCATCTTACCTTTACACAGGAG 3274
QY 1896 -----SerProProValCysGluProValLysCysSerProGluAsnLeuAsn 1913
DB 3275 ATGATGTCACCGTCTTACTACGGAGCACAATTGCGAGTGAATTGAGAAATTGAAT-- 3331
QY 1914 GlyLysTyrIleLeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThr 1933
DB 3332 -----ATTGTGAAGAAATTTTCGCTGAATATTCATCTACATGCTTGGAGCCT 3382
QY 1934 GlyTyrSerLeuGlnGlyProSerIleLeuGluCysThrAlaSerGlyIleTirpAspArg 1953
DB 3383 TCCCTTCAGCGAACAATGAAATACATGTGGCCATTCTGCTGAGATATACGGGATGAT 3442
QY 1954 AlaProPro 1956
DB 3443 GCGAACCAG 3451

RESULT 24

US-08-611-729A-5
Sequence 5, Application US/08611729A
Patent No. 6004924

GENERAL INFORMATION:

APPLICANT: Ish-Horowicz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSER: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,729A
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6464 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 371..4024
US-08-611-729A-5
Alignment Scores:
Pred. No.: 7,41e-52 Length: 6464
Score: 759.00 Matches: 288
Percent Similarity: 34.13% Conservative: 143
Best Local Similarity: 22.80% Mismatches: 402
Query Match: 3.80% Indels: 430
DB: 3 Gaps: 57
US-09-977-053-4 (1-3571) x US-08-611-729A-5 (1-6464)
QY 834 CysSerAspAlaGlu-----AspIleAspCysArgLeuGluAsnLeuThrLys 850
DB 533 TGGCGCGCGCCCGGAACCCCGGAGACCGCAAGTGCACCCCGCAGAGTGTGACACATAC 592
QY 851 ---LysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe---AlaIleGlyPro 868
DB 593 TTCAAAGTGTGCTCAAG-----GAGTATCAGTCCCGGTCCAGCGCGGGGGGCC 643
QY 869 GlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrVal 888
DB 644 TGCAGCTTCGGCTCAGGG----- 661
QY 889 GlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaPro 908
DB 662 -----TCCAGCGCTGTCATCGGGGCAACACCTTCAACCTCAAGGCCAGCGCGCGC 712
QY 909 LeuSerAspTyrLysIleLysLeuIlePheAsnIle-----ThrAla 922
DB 713 AACGCCCGAACCGCATCGTCTGCTTTCAGTTTCGCTGCGCGAGTCTCTATACGTTG 772
QY 923 SerValProLeuProAspGluArgAsnAspThrLeuGluTrpGlu----- 937
DB 773 CTGTGGAGGCGTGGGATTCAGTAATAGACACCGTCAACCTGACAGTATTATTGAAAG 832
QY 938 -----AsnGlnArgLeuLeuGlnThrLeuGluThrIleThr 950
DB 833 GTTCTCAGTGGGCGATGATCAACCCCGCGGAGTGGCAGACGCTGAACAGACAGACG 892
QY 951 AsnLysLeuLys-----ArgThrLeuAsnLysAspProMetTyrSerPhe 965
DB 893 GCGCTTCCCGCCCTTGTAGTATCAGATCGCGGTGACCTGTGATGACTACTACTATGCTTT 952
QY 966 GlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGluThrLysLysAlaSer 985
DB 953 -----GGCTGTAAT 961
QY 986 ProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGly 1005
DB 962 AGTTCTCGCGCCCGCAGAGATGACTTCTTTGGA----- 994
QY 1006 ThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAsp 1025
DB 995 -----CACTATGCTGTGAC-----CAGAAATGGCAACAAACTTGC 1030
QY 1026 GluGluGlyGlnLeu-----GluCysLys-----LeuCysProSerGlyMetTyrThr 1041
DB 1031 ATGGAAGGCTGGATGGGCGCCCGGAATGTAAACAGAGCTATTTGCCGACAGGTGCGAGTCT 1090
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 1091 AAGCATGGGTCTTGCAAACTCCCGAGGTGACTGCGAGGTGCCAG-----TAC 1135
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCys-----ProLeuGlyThrTyr--- 1077
DB 1136 GGCTGGCAAGGCTG---TACTGTGATAAGTGCATCCACACCGGGGTGCGTCCACGGC 1192
QY 1078 -----GlnProLysPheGlySerArgSerCysLeu 1087
DB 1193 ATCTGTAATGAGCCCTGGCAGTGCCTCTGTGAGACCAACTGGGCGCGCGCTCTGT--- 1249

QY 1088 SerCysProGluAsnThrSerThrValIysArgGlyAlaValAsnIleSerAlaCysGly 1107
DB : : : : :
1250 -----GACAAAGATCTCAATTACTGTGGG 1273
QY 1108 Val -----ProCysProGluGly -----LysPheSer 1116
DB : : : : :
1274 ACTCATCAGCCGTCTCAACCGGGAACTTTAGCAACACAGCCCTGACAAATATCAG 1333
QY 1117 ArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGly 1136
DB : : : : :
1334 TGTTC-----TGCCCTGAGGGGTATTCAGGACCCCAACTGT--- 1369
QY 1137 LysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSer 1156
DB : : : : :
1369 ----- 1369
QY 1157 IleThrGluCysSerSerPheSerThrPheSerAlaAlaGluSerValPro 1176
DB : : : : :
1369 ----- 1369
QY 1177 ProAlaSerLeuGlyHisIleLysIysArgHisGluIleSerSerGlnValPheHisGlu 1196
DB : : : : :
1370 -----GAAATTGCTGAG-----CAGGCC 1387
QY 1197 CysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrVal 1216
DB : : : : :
1388 TGCCCTCTGTATCCCTGTCAACACAGGACGCTGTAAGGAGACCTCCCTGGGCTTTGAG 1447
QY 1217 CysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSer 1236
DB : : : : :
1448 TGTGAGTGTCCCGAGCTGGACCGCCGCCACATGCTCTACAACTGATGACTGTCT 1507
QY 1237 ProLeuProCysLeuAsnAsnGlyValCysIysAspLeuValGlyLeuPheIleCysGlu 1256
DB : : : : :
1508 CTAATAACTGTTCACACGGGGCCACCTGCCAGGACCTGTTAAACGATTAAAGTGTG 1567
QY 1257 CysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSer 1276
DB : : : : :
1568 TGCCCCCACCACAGTGGGAAACCGTCCAGTTAGATGCAAAATGAATGTGAGGCCAAA 1627
QY 1277 ProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysVal 1296
DB : : : : :
1628 CTTGTGTAAACGCCAAATCTGTGAAGATCTCATGTCAGCTACTACTGCACTGTCTT 1687
QY 1297 LysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCys 1316
DB : : : : :
1688 CCGGCTGGTGGTCAAGATTGTGACATAAATATTATGACTGCTTGGCCAG---TGT 1744
QY 1317 LeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysIysCysProProGly 1336
DB : : : : :
1745 CAGAAATGACGCTCTCTGTCGGATTGTTAATGTTATCGCTGTATCTGTCCACTGTGC 1804
QY 1337 PheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsn 1356
DB : : : : :
1805 TATGACGGCATCACTGTGAGAGACATCGATGATGTGCCAGAACCCCTGTTTGAAT 1864
QY 1357 GlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThr 1376
DB : : : : :
1865 GGGGTCACTGTCAAGATGAATCAACAGATTCAGTGTCTGTCTCCACTGTTCTCT 1924
QY 1377 GlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAla 1396
DB : : : : :
1925 GGAACCTCTGTGACGTGACATCGATATTATGTGAGCCCTAATCCCTGCCAGAACCGTGC 1984
QY 1397 ThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGln 1416
DB : : : : :
1985 CAGTCTACACCGCTGCCAGTCACTATTCTTCAGAGTGCCTCCGAGACTATCAGGGCAG 2044
QY 1417 ArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyr 1436
DB : : : : :
2045 AACTGCTCACCTGAAAGACCACCTGCCGACGACCCCTGTGAAGTG----- 2092

QY 1437 GlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyr 1456
DB : : : : :
2093 -----ATTGAC-----AGCTGCACAGTGGCC 2113
QY 1457 MetLysSerSerAspAspMetAsnTyrGlyThrPro-----IleSerTyrAlaValAsp 1474
DB : : : : :
2114 ATGGCTTCCACGAC-----ACACCTGAAGGGGTGCGGTATATTCTCTCC 2158
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuValAsn 1494
DB : : : : :
2159 AAC-----GTCTGTGTCTCTCCAC 2176
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAla 1514
DB : : : : :
2177 GGGAG-----TGCAGAGTCTAGTCGGAGGCAATTC----- 2209
QY 1515 IleThrThrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
DB : : : : :
2210 ---ACCTGTGACTGTAAACAAAGGCTTCACGGGAACATACTGCCATGAAATATTATGAC 2266
QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGlyGlyAlaLeuValLeuGly 1554
DB : : : : :
2267 TGTGAGAGCAAC-----CCTGTGAAACCGTGGCACTTGCATCGATGGT 2311
QY 1555 GlnGluGlnAspLysGlyGluGlyPheSerProAlaGluSerPheValGlySerIle 1574
DB : : : : :
2312 GTCAACTCTCTACAG-----TGCACTGTGTAGTACCGCTGGGAGGGGCTAC 2359
QY 1575 SerGlnLeuAsnLeuTyrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThr 1594
DB : : : : :
2360 TGTGAACCAATATTATGACTGCAGCCAGAACCCC----- 2395
QY 1595 SerCysProGluGluLeuSerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGly 1614
DB : : : : :
2396 ---TGCCACAAT----- 2404
QY 1615 IleValGlyLysValLysIleAspSerLysSerIlePheCysSerAspCysProArg--- 1633
DB : : : : :
2405 -----GGGGCAGCTGTCCGACCTGCTGTCATGACTTCTACTGTGACTGTAAATGGG 2458
QY 1634 LeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLysProGlySerLys 1653
DB : : : : :
2459 TGGAAAGGAAGAACCTGCCACTCAGTGCAGTGTGATGAGGCCACGTCGCAACAC 2518
QY 1654 ValAsnLeuPheCysAspProGly-----PheGlnLeuValGlyAsnProValGlnTyr 1671
DB : : : : :
2519 GGTGGCACCCTGTATGATGAGGGGGATGCTTTTAAAG----- 2554
QY 1672 CysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSerCysGly 1691
DB : : : : :
2555 TGCATGTCTCTGGCGCTGGGAGGAACAACCTGTAAACATAGCCCGAACAAGTAGCTGC 2614
QY 1692 ValProProLeuGluAsnGlyPheHisSerAlaAspPheTyrAlaGlySerThr 1711
DB : : : : :
2615 CTGCCCAACCCCTGCCATAATGGG----- 2638
QY 1712 ValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThr 1731
DB : : : : :
2639 -----GGCAGATGTGTGTCAACGGCGAGTCTTTTACGTGGTC 2677
QY 1732 AspAsnGlySerTyrAsnGlyValSerProSerCysLeuAspValAspGluCysAlaVal 1751
DB : : : : :
2678 TCCAGGAAGCTGGAGGG-----CCCATCTGTGTCTCAGATATACC----- 2719
QY 1752 GlySerAspCysSerGluHis-----AlaSerCysLeuAsnValAspGly 1766
DB : : : : :
2720 ---AATGACTGAGCCCTCATCCCTGTTACAAACAGCGCACCTGTGTGATGGAGACAAC 2776
QY 1767 SerTyrIleCysSerCysValProProTyrThrGlyAsp----- 1779
DB : : : : :
2777 TGGTACCGGTGAATGTCCCGGGTTTGTGGGCCCGACGTGCAGAAATAAACATCAAT 2836
QY 1780 -----GlyLysAsnCysAlaGluProfile----- 1787

```
Db 2837 GAATGCCAGTCTTCACTTGTGCTTGGAGCCACCTGTGTGATGAGATCAATGCTAC 2896
Qy 1788 LysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluLeuIleThrVal 1807
Db 2897 CGGTGTGTGCTGCTT-----CCAGGGCACAGTGTGCTCCAG-----2932
Qy 1808 GlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysIle 1827
Db 2933 -----TGCCAGGAA-----GTTTCAGGGAGA 2953
Qy 1828 ThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro-----1840
Db 2954 CTTGATCACCATGGGGAGT-----GTGATACCAGATGGGCCAAATGGGATGAT 3004
Qy 1841 -----TyrCysLysAlaValSerCysGly 1848
Db 3005 GACTGTATACCTGCCAGTGCCTGATGAGCGATGCTGCTCAAAAGGTCTGTGTGGC 3064
Qy 1849 ---LysProAlaIleProGluAsnGly-----CysIle 1858
Db 3065 CTTGACCTTGTGCTGTCTCCAAAGGGCCAGAGTGGCCCGGCGAGAGCTGCATC 3124
Qy 1859 -----GluGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLys 1875
Db 3125 CCATCTCTGGAGCAGCAGCTGCTTC-----GTCCACCCCTGC-----3160
Qy 1876 GlyTyrThrLeuAlaGlyAspLysGluSerSerCysLeu-----1888
Db 3161 -----ACTGGTGTGGCGAGTGTGCTTCCAGTCTCCAGCGCGGTGAAGACAAAGTGC 3214
Qy 1889 -----AlaAsnSerSerTrpSerHis-----1895
Db 3215 ACCTCTGACTCTTATACAGGAGTAATCTGCGACATACATCATTTACCTTTAACAAGGAG 3274
Qy 1896 -----SerProValCysGluProValLysCysSerSerProGluAsnIleAsnAsn 1913
Db 3275 ATGATGTACACAGGTCTTACTACGGAGCACATTTGCGAGTGAATTGAGGATTTGAT---3331
Qy 1914 GlyLysTyrIleLeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThr 1933
Db 3332 -----ATTTGAAGAATGTTTCCGCTGAATATTTCAATCTACATCGCTTCGAGCCT 3382
Qy 1934 GlyTyrSerLeuGlnGlyProSerIleIleGluCysThrAlaSerGlyIleTrpAspArg 1953
Db 3383 TCCCTTTCAGCGAACAATGAATATACATGATGCGCAATTTCTGCTGAAGATATACGGGATGAT 3442
Qy 1954 AlaProPro 1956
Db 3443 GGGNACCCG 3451
```

RESULT 25

```
PCT-US91-05059-1
; Sequence 1, Application PC/TUS9105059
; GENERAL INFORMATION:
; APPLICANT: Regents of the Board of the, University of
; APPLICANT: Oklahoma
; TITLE OF INVENTION: Functionally Active Selectin-Derived
; TITLE OF INVENTION: Peptides
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Kilpatrick & Cody
; STREET: 100 Peachtree Street, Suite 3100
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: US
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/05059
; FILING DATE: 19910717
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/320408
; FILING DATE: 08-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/554199
; FILING DATE: 17-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF110CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3144 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Blood
; CELL TYPE: Endothelial
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 2833..2838
; OTHER INFORMATION: /note= "Potential polyadenylation
; OTHER INFORMATION: signals"
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 3124..3130
; OTHER INFORMATION: /note= "Potential polyadenylation
; OTHER INFORMATION: signal"
; PCT-US91-05059-1
;
; Alignment Scores:
; Pred. No.: 2,26e-52 Length: 3144
; Score: 758.50 Matches: 272
; Percent Similarity: 33.42% Conservative: 124
; Best Local Similarity: 22.95% Mismatches: 352
; Query Match: 3.80% Indels: 438
; DB: 5 Gaps: 49
;
; US-09-977-053-4 (1-3571) x PCT-US91-05059-1 (1-3144)
;
; Qy 1734 GlySerTrpAsnGly-----ValSerProSer 1742
; Db 465 GGCAGTGAATGATGAGCAGCTCTTGAAGAAAAGCAGCATTTGTTACAGACCTCC 524
; Qy 1743 CysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeu 1762
; Db 525 TGCCAGGACATG-----TCTGACGACAAACAGAGAGTGCCTC 563
; Qy 1763 AsnValAspGlySerTyrIleCysSerCysValProTyrThrGlyAspGly-----1780
; Db 564 GAGACCATCGGAACCTACACCTGCTCTCTGTTACCTGGATTCTATGGCCAGAAATGTGA 623
; Qy 1781 -----LysAsnCysAlaGlu 1785
; Db 624 TAGCTGAGAGAGTGTGGAGAACTTGAGCTCCCTCAACACGCTGCTCATGAACCTGACGCCAC 683
; Qy 1786 ProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGluIleThr 1805
; Db 684 CTTCTG-----GGAAAC-----TTC 698
; Qy 1806 ThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThr 1825
```

Db 699 TCTTTAACTCCGAGTCTCCACTGCAGCGGTACCAAGTAAATGGGCCAGC 758
Qy 1826 LysIleThrCysLeuGluSerGlyGluThrPheHisLeuIleProTyrCysLysAlaVal 1845
Db 759 AAGCTGGAATGCTGGCTTCTGGAATCTGGACAATAAGCCCTCCACAGTGTATTAGTCC 818
Qy 1846 SerCysGlyLysProAlaIleProGluAsnGlyCysIleGluLeuAlaPheThrPhe 1865
Db 819 CAGTCCCCACCCTCGAAGATCTCTGACGAGA 851
Qy 1866 GlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLysGluSer 1885
Db 852 -----AACATG 857
Qy 1886 SerCysLeuAlaAsnSerTrpSerHisSerProProValCysGluProValLysCys 1905
Db 858 ATCTGCTT -----CATCT 872
Qy 1906 SerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeuThrTyrLeuSer 1925
Db 873 -----GCAAAAGCAATCCAGCATCAGTCT 896
Qy 1926 ThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIleGluCys 1945
Db 897 AGCTGCAGCTTCAGTGTGAGAGGAGTTTGCATTGTGGACCCGAAAGTGTGCAATGC 956
Qy 1946 ThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCysGly 1964
Db 957 ACAGCTCGGGGTATGACAGCCCGCCAGCGTGTAAAGCTGTGCAGTGTGCAGCAC 1016
Qy 1965 ---GluProProAla-----IleLysAspAlaValIleThrGlyAsnAsnPheThrPhe 1981
Db 1017 CTGGAAGCCCCAGTGAAGGAACCACTGAGTGTTCATCCGCTCAGCTCTTTGCCAT 1076
Qy 1982 ArgAsnThrValThrTyrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIle 2001
Db 1077 GCCTCCAGCTGCAATTTGAGTGCCAGCCGCTACAGAGTGAGGGGCTTGACATGCTC 1136
Qy 2002 GluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCys 2021
Db 1137 CGCTGCAATGACTCGGACACTGGTCTGCACCTTGCCAACTGTGAGGCTATTTCGTGT 1196
Qy 2022 Asp-----GluProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhe 2039
Db 1197 GAGCGCTGGAGAGTCTGTC-----CACGGAAGCATGGAT----- 1232
Qy 2040 GlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeu 2059
Db 1233 -----TGCTCT----- 1238
Qy 2060 LeuCysAsnAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIle 2079
Db 1238 ----- 1238
Qy 2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLys 2099
Db 1239 -----CCATCTCTG----- 1247
Qy 2100 AlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsn 2119
Db 1248 AGAGCGTTTCAGTATGACACCACTGATGCTTCCGCTGTGCTGGAAGGTTTCATGCTGAGA 1307
Qy 2120 ThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsn---ProSerProMetSer 2138
Db 1308 GGAGCCGATATAGTTGGTGTGATAACTTGGGACAGTGGACAGCACCAGCCGCCAGTC--- 1364
Qy 2139 IleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSer 2158
Db 1365 -----TGCAAGCTTTTGAGTGCAGGATCTCCCA---GTTCCAAATGAGCCCGGGTG 1415
Qy 2159 GlySerAsnTyrSerPheGlyAla-----MetValAlaTyrSerCysAsn 2173
Db 2173 ----- 2173

Db 1416 AACTGCTCCACCCCTTGGTGCCTTTAGGTACCAAGTCAGTCTGCAGCTTCACCTGCAAT 1475
Qy 2174 LysGlyPheTyrIleLysGlyGluLysSerThrCysGluAlaThrGlyGlnTrpSer 2193
Db 1476 GAAGCTTTCCTCTGCTGGAGCAAGTGTGTACAGTGTCTGGCTACTGGAACCTGGAAT 1535
Qy 2194 SerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsnGly 2213
Db 1536 TCTGTTCTCCAGAAATCCAGCCATTCCCTGCACACCTTGTGCTAAGCCCTCAGAATGA 1595
Qy 2214 PheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsnPro 2233
Db 1595 ----- 1595
Qy 2234 GlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHisSer 2253
Db 1595 ----- 1595
Qy 2254 GluSerProLeuMetCysValProLeuAspCysGlyLysProProProIleGlnAsnGly 2273
Db 1596 -----ACAATGACCTGTGTT-----CAACCTCTTGGAACTTCC 1628
Qy 2274 PheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGly 2293
Db 1629 AGTTATAAA-----TCCACATGTCAATTTCATCTGTGACGAGGGA 1667
Qy 2294 TyrGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLys 2313
Db 1668 TATTCTTTGTCTGGACCAAGAAAGATTGATGTATCGATCGGACCTGGACA---GAC 1724
Qy 2314 SerAsnProLysCysMetProAlaLysCysProGlu-----ProProLeuLeuGlu 2330
Db 1725 TCCCAACCAATGTGTGAAGCCATCAGTGTCCAGAACTCTTTGCCCCAGAG---CAG 1778
Qy 2331 AsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThr-----Phe 2347
Db 1779 GCGAGCTGGATTGTCTGACACTCGTGAGAAATCAATGTGGTCTCCACCTGTCTCATTT 1838
Qy 2348 SerCysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGln 2367
Db 1839 TCTTGTAAACAATGGCTTTTAAGCTGGAGGGGCCAATAATGTGGAATGCACTTCTGGA 1898
Qy 2368 GlnTrpAsnAspSerPheProValCysLysIleValLeuCysThrProProLeuIle 2387
Db 1899 AGATGCTCAGTACTCCACCACTGCAAGCATAGCATCATCTTCTTCTTCCA----- 1952
Qy 2388 SerPheGlyValProIleProSer----- 2395
Db 1953 -----GGGTTGCAATGTCCAGCCCTCACCCTCTCGGCGAGGAAACCATGTACTGTAGG 2006
Qy 2396 -----SerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGlyPhe 2412
Db 2007 CATCATCCGGGAACCTTTGGTTTAAATACCACTGTGTACTTTGGCTGCAACCGCTGGATTC 2066
Qy 2413 PheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeu 2432
Db 2067 ACACATATAGGAGACAGCACTCTCAGCTGCACACCTTCAAGCAATGAGCAGCACTAACT 2126
Qy 2433 ProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIleIleAsp 2452
Db 2127 CCAGCATGCAGAGCTGTGAATGCTCAGAA----- 2156
Qy 2453 ValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeu 2472
Db 2157 -----CTACATGTTAATAAGCCA-----ATA 2177
Qy 2473 ValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThr 2492
Db 2178 GCGATGAACCTGCTCCAACTC-----TGG----- 2201
Qy 2493 CysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThr 2512
Db 2202 -----GGAAACTTTCAGT----- 2213

Search completed: May 11, 2004, 22:16:16
Job time : 1495.23 secs

```
QY 2513 AsnLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGly 2532
Db |||||
Db -----TATGATCAATCTGCTCTTCCATTGCTAGAGGGCCAGTTTACTTAAATGGC 2264
QY 2533 ProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsn 2552
Db |||||
Db TCTGCACAAACAGCATGCCAAGAGAAATGGCCACTGGTCAACTACCGTGCCCAACCTGCCAA 2324
QY 2553 AlaIleHisCysAspSerProGlnProIleGlu-----AsnGlyPheVal 2567
Db |||||
Db GCAGG-ACCATTGACTATCCAGGAGCCCTGACTTCTGGTGAGCGGTGGCTTCTAC 2393
QY 2568 GluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGln 2587
Db |||||
Db AATAGGTCTGATAATGGGTGGAGCT-----CCTGGCTTT--- 2419
QY 2588 ValAlaGlyHisAlaMetGlnThrCysGluSerGlyTrpSerSerIleProThr 2607
Db |||||
Db ---GCTAAGAAAGCGTTTTCAGACA-----AAAGATGATGGGAA----- 2455
QY 2608 CysMetProIleAspCysGlyLeuProProHis----- 2618
Db |||||
Db ATGCCCTTGAATCTCTCAGCCACCTAGGAACATATGGAGTTTTCACAAACGCTGC 2512
QY 2618 ----- 2618
Db 2513 ATTTGACCCGAGTCCTTAAGSTTTCATAAACCCCATGAATCAAGACATGGAATTACC 2572
QY 2619 -----IleAsp-----Phe 2621
Db |||||
Db TTAGATTAGCTCTGGACACGCTGTGGACCCGCTCTGGACCAACCCCTGTTTCTCTGAGTT 2632
QY 2622 GlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMetMet 2641
Db |||||
Db TGGATTGTGGTACAATCTCAATCTCAACCT----- 2665
QY 2642 GluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTyr 2661
Db |||||
Db -----ACCACCCCTCTCTGTCCC----- 2683
QY 2662 GluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetVal 2691
Db |||||
Db -----ACCTCTCTCTCTCTCTGTAACACCAAGCCACAGAGCCAGGACAAATGTTCTGTC 2737
QY 2682 SerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGlnGlu 2701
Db |||||
Db AGTAGTCTGTGCTTTGACTCACCTGTTACTTGAATATACAGTG-----AACCAAGA 2791
QY 2702 AspGlyThrTrpAsn-----GlySerAlaProSerCysIleSerIleGluCysAsp 2718
Db |||||
Db GAC-----TGGAGCATCTGACTCACAAAGAACCCAGACTGTGGAGAAATAAATAATACCT 2845
QY 2719 LeuProThrAlaProGlu--AsnGlyPheLeuArgPheThrGluThrSerMetGlySerAl 2738
Db |||||
Db CTTTATTTTGTGATGAAGGAGGTTTCTCCACTTTGTTGGAAGCAGGTGGCATCTCT 2905
QY 2738 aVal---GlnTyrSerCysLys--ProGlyHisIleLeuAlaGlySerAspLeuArgLeu 2757
Db |||||
Db AATTGGAAGAAATCTCTGTAGCATCTCTGTGAGTCTCCAGTGGT-----T 2950
QY 2757 Ys-LeuGluAsnArgLysTrpSerGlyAla-----SerProArgCysGlu 2771
Db |||||
Db GCTGTTGATGAGGCTCTTGGACCTCTGCTCTGAGGCTTCCAGAGAGTCTCTCTGATGGC 3010
QY 2772 AlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyr 2791
Db |||||
Db ACCAGAGGCTGCAGAGAGCCCAAGAAATCAAGCTAGAGGCCACATG-----TCACCGTGG 3064
QY 2792 ThrTyrLeu 2794
Db |||||
Db ACCTTCTCTG 3073
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2004, 07:07:47 ; Search time 2398.04 Seconds
(without alignments)
6326.134 Million cell updates/sec

Title: US-09-977-053-4

Perfect score: 19973

Sequence: 1 MWPLRAFCWGLALVSGWAT.....CHCLSSWTGHCNCRKRTGCF 3571

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO spool_p/US09977053/runat_06052004_075942_18178/app_query.fasta_1.5710
-DB=N Geneseq 26Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=25
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09977053 @CGN 1.1 4042 @runat_06052004_075942_18178 -NCPU=6 -ICPU=3
-NO MMAP -LARGEBQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_29Jan04:*
1: geneseqm1980s:*
2: geneseqm1990s:*
3: geneseqm2000s:*
4: geneseqm2001as:*
5: geneseqm2001bs:*
6: geneseqm2002s:*
7: geneseqm2003as:*
8: geneseqm2003bs:*
9: geneseqm2003cs:*
10: geneseqm2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19958	99.9	10878	6	AAD32025 Human C3b
2	19905.5	99.7	11546	8	AAD57253 Human C3b
3	19616	98.2	11152	6	ABT08491 Human nov
4	19566	98.0	11158	6	ABT08490 Human nov
5	16707	83.6	11230	6	AAD32026 Human C3b
6	9400	47.1	5421	4	AAK94920 Human ful
7	8427	42.2	6153	6	ABA03880 Human POL
8	8400.5	42.1	6152	7	ABX56476 cDNA enco

9	8391	42.0	6153	9	AAD58963	Aad58963 Human com
10	7537	37.7	5124	4	AAH16667	Aah16667 Human cDN
11	7471	37.4	3921	6	ABK13101	Abk13101 Human REP
12	6950	34.8	4088	4	AAK94919	Aak94919 Human ful
13	6452	32.3	3804	4	AAS28968	Aas28968 cDNA enco
14	6452	32.3	3804	4	AAS26868	Aas26868 Human cDN
15	6452	32.3	3804	4	ABA06548	Aba06548 Human cDN
16	6452	32.3	3804	4	AAS31587	Aas31587 cDNA enco
17	6452	32.3	3804	4	ABK43685	Abk43685 DNA enco
18	6452	32.3	3804	6	ABT07803	Abt07803 Novel hum
19	6452	32.3	3804	6	ABV83885	Abv83885 Human pol
20	6044	31.7	3696	4	AAH75787	Aah75787 Receptor
21	6044	31.7	3706	4	ABN93421	Abn93421 Human gen
22	5485	27.5	3448	6	AB211152	Ab211152 Human pol
23	5012	25.1	2929	4	AAS03887	Aas03887 Human sec
24	4971	24.9	2637	7	ABX34482	Abx34482 Human mdd
25	4930	24.7	3262	6	AB511807	Ab511807 Human mdd
26	4730	23.7	3128	6	ABK54124	Abk54124 cDNA enco
27	4674	23.4	2575	6	ABQ72635	Abq72635 Human MDD
28	4058	20.3	2437	6	ABK94974	Abk94974 Human nov
29	4045	20.3	2444	6	ABK94928	Abk94928 Human nov
30	4039.5	20.2	2450	6	ABT10760	Abt10760 Human bre
31	4039.5	20.2	2450	9	ABE73131	Abe73131 Human cel
32	4026.5	20.2	2539	3	AACT77331	Aac777331 Human ORF
33	3596	18.0	1932	6	ABA03879	Aba03879 Human POL
34	3580	17.9	1952	7	ABX56475	Abx56475 cDNA enco
35	3547	17.8	1952	6	AAD58962	Aad58962 Human com
36	3133	15.7	2084	9	ABK35664	Abk35664 cDNA sequ
37	3073	15.4	1969	4	AAH16567	Aah16567 Human cDN
38	2892	14.5	1880	4	AAH48345	Aah48345 Human gra
39	2831	14.2	1765	4	AAK94846	Aak94846 Human ful
40	2758	13.8	1493	4	AAS28991	Aas28991 cDNA enco
41	2758	13.8	1493	4	ABX26940	Abx26940 Human cDN
42	2758	13.8	1493	4	ABA06730	Aba06730 Human cDN
43	2758	13.8	1493	4	AAS31645	Abk31645 cDNA enco
44	2758	13.8	1493	4	ABK43966	Abk43966 DNA enco
45	2758	13.8	1493	6	ABT07826	Abt07826 Novel hum
46	2758	13.8	1493	6	ABV84067	Abv84067 Human pol
47	2638	13.2	1652	9	ADC35133	Adc35133 Human bre
48	2622	13.1	1747	2	AZA24234	Aza24234 Human nor
49	2506	12.5	1357	6	ABQ72495	Abq72495 Human MDD
50	2479	12.4	1696	6	ABK54146	Abk54146 cDNA enco
51	2374	11.9	1760	3	AAS294659	Aas294659 Human EGF
52	2305	11.5	1709	6	ABA03878	Aba03878 Human POL
53	2305	11.5	1709	7	ABX56474	Abx56474 cDNA enco
54	2290	11.5	1709	9	AAD58961	Aad58961 Human com
55	2284	11.4	1766	4	ABN93419	Abn93419 Human gen
56	2248	11.3	1377	4	ABN93416	Abn93416 Human gen
57	2247	11.3	1640	2	AAT12162	Aat12162 Partial p
58	2037	10.2	9038	5	AAS64290	Aas64290 DNA enco
59	2032	10.2	7313	4	AAI58380	Aai58380 Human pol
60	2032	10.2	7313	8	ADB48349	Adb48349 Novel hum
61	2032	10.2	7385	7	ABX34686	Abx34686 Human mdd
62	2032	10.2	7821	5	AAS64474	Aas64474 DNA enco
63	2027	10.1	6951	1	AAH91477	Aah91477 CR1 prote
64	2027	10.1	6951	2	AAS38150	Aas38150 Human C3b
65	2027	10.1	6951	6	ABK84738	Abk84738 Human cDN
66	2027	10.1	6951	6	ABA91636	Aba91636 Human C3b
67	2027	10.1	6951	9	ADB85074	Adb85074 Farnesyl
68	2016	10.1	6951	2	AAQ11642	Aaq11642 Entire hu
69	2006	10.0	7028	4	ABA09026	Aba09026 Human CR1
70	2006	10.0	7028	4	AAI60166	Aai60166 Human pol
71	1977	9.9	1408	6	ABK35663	Abk35663 cDNA sequ
72	1977	9.9	6951	2	AAQ41867	Aaq41867 CR1 codin
73	1852	9.3	12525	6	AAD33319	Aad33319 Rat C3b/C
74	1825.5	9.1	10433	6	AAD33320	Aad33320 Human C3b
75	1825.5	9.1	10673	6	AAD33318	Aad33318 Human C3b
76	1822	9.1	10136	6	AB564375	Ab564375 Human cub
77	1606.5	8.0	5420	2	AAS38151	Aas38151 Human C3b
78	1606.5	8.0	5420	6	ABL65862	Ab165862 Lung canc
79	1606.5	8.0	5420	6	ABA91637	Aba91637 Human C3b
80	1604.5	8.0	5420	2	AAQ11643	Aaq11643 Partial h
81	1498	7.5	1892	4	AAS26857	Aas26857 Human cDN

82 1462.5 7.3 10489 4 ABL14889
 83 1447 7.2 765 4 AAB08402
 84 1443.5 7.2 13484 4 ABL14888
 85 1379 6.9 3896 4 AAF87127
 86 1349 6.8 3905 4 AAF87126
 87 1341 6.7 706 7 ACAS7318
 88 1304 6.5 801 4 AAB07752
 89 1278 6.4 680 4 AAS28962
 90 1278 6.4 680 4 AAS26930
 91 1278 6.4 680 4 ABA06507
 92 1278 6.4 680 4 AAS31580
 93 1278 6.4 680 6 ABO7797
 94 1278 6.4 680 6 ABO7797
 95 1225.5 6.1 771 4 AAK93472
 96 1217 6.1 3327 2 AAO10989
 97 1217 6.1 4094 9 AAD67537
 98 1182 5.9 8010 6 AAS64376
 99 1173.5 5.9 728 4 AAK93506
 100 1173.5 5.9 728 4 AAK92246

ALIGNMENTS

RESULT 1

AAD32025

ID AAD32025 standard; cDNA; 10878 BP.

XX AC AAD32025;

XX DT 18-JUN-2002 (first entry)

XX DE Human C3b/C4b complement receptor-like cDNA.

XX KW Human; C3b/C4b complement receptor-like protein; CR-like; diabetes;

XX KW immune system disorder; rheumatoid arthritis; psoriatic arthritis;

XX KW inflammatory arthritis; osteoarthritis; inflammatory joint disease;

XX KW autoimmune disease; multiple sclerosis; inflammatory bowel disease;

XX KW transplant rejection; graft versus host disease; atherosclerosis; lupus;

XX KW stroke; Alzheimer's disease; ischemic condition; noctropia; restenosis;

XX KW myocardial infarction; ischemia; metabolic disorder; obesity;

XX KW reproductive disorder; infertility; nervous system disorder;

XX KW gene therapy; immunomodulator; antipsoriatic; antiinflammatory;

XX KW neuroprotective; vasodilator; cardiant; anorectic; gene; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 11..10726

XX FT /*tag= a

XX FT /product= "Human C3b/C4b CR-like protein"

XX WO200210388-A2.

XX PD 07-FEB-2002.

XX PF 24-JUL-2001; 2001WO-US023548.

XX PR 01-AUG-2000; 2000US-0222438P.

XX PA (AMGE-) AMGEN INC.

XX PI Welcher AA, Elliot GS;

XX DR WPI; 2002-257381/30.

XX DR P-PSDB; AAE20146.

XX PT Nucleic acid encoding a novel C3b/C4b Complement Receptor (CR)-like

XX PT nucleic acid molecule, useful for treating, preventing and diagnosing

XX PT rheumatoid arthritis, psoriatic arthritis, inflammatory arthritis, and

XX PT multiple sclerosis.

XX PS Claim 1; Fig 1A-1H; 201pp; English.

XX PS

XX

CC The invention relates to C3b/C4b complement receptor (CR)-like protein
 CC and its corresponding nucleic acid sequence. The C3b/C4b CR-like
 CC polypeptide and nucleic acid molecules may be used to treat, prevent,
 CC ameliorate, diagnose and/or detect diseases such as immune system
 CC disorders such as rheumatoid arthritis, psoriatic arthritis, inflammatory
 CC arthritis, osteoarthritis, inflammatory joint disease, autoimmune
 CC disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease,
 CC transplant rejection, graft versus host disease, nervous system disorders
 CC (e.g. stroke, Alzheimer's disease), ischemic conditions (e.g.
 CC atherosclerosis, restenosis, myocardial infarction, and ischaemia),
 CC metabolic disorders (e.g. obesity and diabetes); and reproductive
 CC disorders and infertility. The C3b/C4b CR-like nucleic acid molecules are
 CC used in gene therapy. The present sequence is human C3b/C4b CR-like cDNA
 XX

SQ Sequence 10878 BP; 2864 A; 2561 C; 2684 G; 2769 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 10878
 Score: 19958.00 Matches: 3566
 Percent Similarity: 99.94% Conservative: 3
 Best Local Similarity: 99.86% Mismatches: 2
 Query Match: 99.92% Indels: 0
 DB: 6 Gaps: 0

US-09-977-053-4 (1-3571) x AAD32025 (1-10878)

QY 1 MetTTPProArgLeuAlaPheCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr 20
 DB 11 ATGTGGCCCTGGCTGGCCCTTTTGTGTGGGTCTGGGGCTGGTTTGGGGTGGGGAC 70
 QY 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
 DB 71 TTTTACGACATGTCTCCCGTGGGCAATTTTTCAGCTTCTTCCCGGAGACGGCGCC 130
 QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyValSerGlyValAlaGlySerArg 60
 DB 131 GGGGGCCCCGGGAGTATCCCGCGCCCGCTCTTGGGACGAGCGGGGGGAGGAGCA 190
 QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
 DB 191 GTGGAGCGCTGGCGCCAGCGTTCCGGCGACGGCTGGCGGTCTGGCGGAGCTCAGCGAG 250
 QY 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg 100
 DB 251 CCCTCTGGAGCTTGTCTTCTGTGGTGGATGATTCCTCAGCGTGGCGGAGTCAACTTCCGC 310
 QY 101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
 DB 311 AGCGAGCTCATGTTCGTCCGAGCTGTCTGTCCGACTTCCCGTGGTGGCGGCGCACG 370
 QY 121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
 DB 371 CGCGTGGCCATCGTGCACCTTCTCGTCCAGAACTACGTGTGTGGTGGCGGCGTTCATC 430
 QY 141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIle 160
 DB 431 TCCACCGCGCGCGCGCGCAGCACAGTGGCGCTCTCTCTCCAGAGATTCCTTGGCTCC 490
 QY 161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeu 180
 DB 491 TCCATCCGAGGTGGCGGCACTTACACCAAGSGCGCTTCCAGCAAGCGCGCAAAATCTT 550
 QY 181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn 200
 DB 551 CTTCATGCTAGAGAAACTCAACAAAAGTTGTATTTCTCATCATCATGATGATTTCAAT 610
 QY 201 GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr 220
 DB 611 GGGGAGACCCCTAGACCAATTCAGCGTCACTCCGAGATTTCAGAGTGGAGATCTTCAC 670
 QY 221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
 DB 671 TTTGGCATATGGCAAGSGAGACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 730

QY 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaLeu 260
DB 731 GAGCACTGTTACCTGCTACACAGTTTGAAGAAATTTGAGGCTTTAGCTGCCCGGCAATG 790
QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
DB 791 CATGAAGATCTACTCTCTGGAGTTTATTCAAGATGATATGCTCCACTGCTCTTATCTT 850
QY 281 CysAspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
DB 851 TGTGATGAAGCAAGCACTGCTGTGACCGAATGGAAGCTGCAAAATGCGGACACACACA 910
QY 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
DB 911 GCCCATTTTGTAGTGCATCTGTGAAAAGGGGTATTACGGGAAGGTCTGCGAGTATGAATGC 970
QY 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
DB 971 ACMGCTTGCCATCGGGGACATACAACTTGAGGCTCACAGAGGATCAGCAGTTGC 1030
QY 341 IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys 360
DB 1031 ATTCCATGCTCTGATGAATAATCACACTCTCCACCTTGGAGCACATCCCCCTGAAGACTGT 1090
QY 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
DB 1091 GTCTGCAGAGAGGATACAGGGCATCTGGCCACAGCTGTGAACTTGTCTCACTGCCCTGCC 1150
QY 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
DB 1151 CTGAACCTCTCCCGAATAATGTTACTTTATTCAAAACACTTGGCAACCACTTCAATGCA 1210
QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLysCys 420
DB 1211 GCCTGTGGGGTCCGATGTCACCTCGGATTTGATCTTGTGGGAAGCAGCATCTTATGT 1270
QY 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
DB 1271 CTACCCAAATGGTTGTGTGCTCCGCTTCAGAGAGTACTTGCAGAGTAAGAACTGTCTCAT 1330
QY 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1331 CTCCCGCAGCCGAAACATGGCCACATCAGCTGTTCTACAGGGAAATGTTATATPAGACA 1390
QY 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1391 ACATGTTTGTGCTGTGATGAAGGTTACAGACTAGAGGCACTGATAGCTTACTTGT 1450
QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1451 CAAGGAACAGCCAGTGGGATGGGCCAGAACCCCGGTGTGTGGAGCGCCACTGTTCACCC 1510
QY 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
DB 1511 TTTCAATGCCCCAAGATATCATCATATCCCCCAACTGTGGCAAGCAGCCAGCCAAA 1570
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1571 TTTGGACGATCTGCTATGATGTAAGTTCGCCCAAGGGTTTCATTTTATCTCGAGTCAAGAA 1630
QY 541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLys 560
DB 1631 ATGCTGAGATGTACCACTTCTGGAAAATGGAATGTGCGAGTTTCAGGCGAGCTGTGTGAAA 1690
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
DB 1691 GACGTGGAGCTCTCTCAATCAACTGTCTTAAAGACATAGAGGCTAAGACTCTGGAACAG 1750
QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 1751 CAAGATTCTGCCAATGTGTACCTGGCAGATTTCCAAACAGCTAAAGACAACTCTGTGTGAAAG 1810

QY 601 ValSerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAspVal 620
DB 1811 GTGTGAGTCCAGCTTATCCAGCTTTCACCCACCTTACCTTTCCCGAGTTGGAGATGTT 1870
QY 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
DB 1871 GCTATCGTATACACGGCACTGACCTATCCGGCAACCAAGCCAGCTGCATTTCCATATC 1930
QY 641 LysValIleAspAlaGluProValIleAspTrpCysArgSerProProValGln 660
DB 1931 AAGGTTATTATGATGAGAACCACTGTTCATAGACTGGTGCAGATCTCCACCTCCCGCTCCAG 1990
QY 661 ValSerGluLysValHisAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
DB 1991 GTCTCGGAGAGGTACATGTCGCCAGCTGGATGAGCTCAGTTCTCAGACACTCAGGG 2050
QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
DB 2051 GCTGAATTTGTCATTTACCAGAGTCTATACACAGGAGACCTTTTCCCTCAAGGGGAGACT 2110
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
DB 2111 ATAGTACAGTATACAGCACTGACCCCTCAGGCAATAACAGCATGTGATATCCATATT 2170
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
DB 2171 GTCATAAAGGTTCTCCCTGTGAATTTCCATTACACCTGTAAATGGGATTTTATATGC 2230
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
DB 2231 ACTCCAGATAATACTCGAGTCACTGATACATTAACTTGTGGAGGGCTATGATTTTCA 2290
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
DB 2291 GAAGGGTCTACTGACAGTATTATTGTGCTTATGAGATGCGCTCTGGAACCAACATAT 2350
QY 781 ThrThrGlnTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
DB 2351 ACCACTGAATGGCGAGACTGTGCCAAAAGCTTTTGTCTAACCAAGGGTTCAAGTCTTT 2410
QY 801 GluMetPheTyrLysAlaLysArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
DB 2411 GAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAGTTTCTGAA 2470
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
DB 2471 GCATTTTGAGACGACCTCGGAAAAATGTTCCCATCATTTTGTAGTGTGAGAGACATT 2530
QY 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
DB 2531 GACTGCGAGACTGGAGGAGAACCTTGACCCAAAATATTTGCTAGAAATATATATGACTAT 2590
QY 861 GluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
DB 2591 GAAAAATGGCTTTGCAATTTGGACCAAGCTGGCTGGCTGAGCTAATAGGCTGATCTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
DB 2651 TACGATGACTTCTCGGACACTGTGCAAGAACAGCCCAAGCATCGGCAATGCGCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
DB 2711 TCACGAAATTAAGAGAGTGGCCCATTTATCTGACTATTAATTAAGTTAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
DB 2771 ACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTTGAATGGGAAAAATCAGCAA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysSer 960
DB 2831 CGACTCTCTCAGACATTTGGAAACTATCACAAATAACTGAAGAGGACTCTCACAAAGAC 2890
QY 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuAlaAspSerAsnSerLeuGlu 980

2891 |||||CCCCATGATTCCTTTTCAGCTTGCATCAGAAATACCTTATAGCCGACAGCAATTCATTAGAA 2950
QY ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
2951 ACACAAAAGGCTTCCCTCTTCGCAGACAGGCTCAGTCTGAGAGGGGTATGTGTCTC 3010
QY AenCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
3011 AATTGCCCTTTGGGAACCTTATATATCTGGAACATTTCCACCTGTGAAAGCTGCCGATC 3070
QY GlySerTyrGluAspGluGlyGlnLeuGluCysLysLeuLeuCysProSerGlyMetTyr 1040
3071 GGATCCTATCAAGATGAAGAAGGCAACTGAGTGCACCTTTGCCCCCTCTGGGATGTAC 3130
QY ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
3131 ACGGAATATATCCATTCAGAAACATCTCTGATTGTAAAGCTCAGTGTAAACAGGSCACC 3190
QY TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
3191 TACTCATACAGTGGACTTGGAGCTTGTGAATCGTGTCCACTGGGCACTTATCAGCCAAA 3250
QY PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
3251 TTGGTTCCTCCGAGCTGCCCTCTCGTGTCCAGAAAACACCTCAACTGTGAAAGAGGAGCC 3310
QY ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
3311 GTGACATTTCTGCATGTGGAGTCTCTGTGCCAGAGGAAATCTCGCGTTCTCGGTTA 3370
QY MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
3371 ATGCCCTGTCCACCATGTCTCGTGAATATACCAACCTAAATGACAGGAGGCGCTTCTGC 3430
QY LeuAlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGluCys 1160
3431 CTGGCCTGTCCCTTTTATGGAATACCCCATTCGCTGGTTCAGATCCATCAGAAATGT 3490
QY SerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu 1180
3491 TCAAGTTTATGTTCAACTTCTCAGCGGCAGAGAGGAGTGTGGTGCCTCTCTT 3550
QY GlyHisIleLysArgHisGluIleSerSerGlnValPheHisGluCysPheAsn 1200
3551 GGACATATTAATAAGAGGCAATGAATCAGCAGTCAGGTTTCCATGAATGCTTTTAAAC 3610
QY ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
3611 CTTTGCACATAGTGGACCTCCAGCAACTTGGGGGTGGTATGTTTGTCTCTGTCCA 3670
QY LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
3671 CTTGGATATACAGGCTTAAAGTGTGAACACAGACATCGATGAGTGCAGCCCATCTGCTGC 3730
QY LeuAsnAsnGlyValCysLysAspLeuValGlyPheIleCysGluCysProSerGly 1260
3731 CTCACAAATGGAGTTTGTAAAGACCTTAGTTGGGGAATTCATTGTGAGTGCCTCAGGT 3790
QY TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsn 1280
3791 TACACAGGTCAGCGGTGTGAAGAAATATATATAGTGTAGTGTCCAGTCTCTGTTTAAAT 3850
QY LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
3851 AAAGGAATCTGTGTGATGGTGTGGCTGCGTATCGTTCACATGTGTGAAGGATTTGTA 3910
QY GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
3911 GGCTTCATTTGTAACAGAAATGAATGTCAGTCAAAACCATGCTTAAATATGCA 3970
QY ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340

3971 GTCTGTGAAGACCAAGTGTGGGGATTCCTTGTGCAAAATGCCACCTCGATTTTGGGTACC 4030
QY ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
4031 CGATGTGGAAAGAAACGTCGATGAGTGTCTCAGTCAGCCCATGCAAAATGAGCTACCTGT 4090
QY LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
4091 AAAGACGGTGCATAGCTTCAGATGCCCTGTGTGCGAGCTGGCTTCACAGGATCACCTGT 4150
QY GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
4151 GAATTGAACTCATGAAATGTCTGCTATTCATGATGAAATCAGGCCACCTGTGTGGAT 4210
QY GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
4211 GAATTAATTCATACAGTTGTAATGTAGCCAGGATTTTCAGGCCAAAGGTGTGAACA 4270
QY GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
4271 GAACAGCTTACAGGCTTTAACTGGATTTTGAAGTTTCTGGCATCTATGGATATATGTATG 4330
QY LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSer 1460
4331 CTAGATGGCATGCTCCATCTCTCCATGCTCTAACCTGTACCTTCTGGATGAAATCTCT 4390
QY AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
4391 GACGACATGAACCTATGAAACACCAATCTCTATGACAGTTGATTAACGCGCAGCAGCAATAC 4450
QY LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
4451 TTGCTCCTCACTGATTTATTAACGCTCGGTCTTTTATGTGAATGGCAGGAAAGATAACA 4510
QY AsnCysProSerValAsnAspGlyArgTyrHisHisIleAlaIleThrTyrThrSerAla 1520
4511 AACTGTCCCTCGTGGATGATGCGAGATGGCATCATATTCATCACTTCGACACAGTGCC 4570
QY AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
4571 AATGGCATCTGGAAAGCTATATATCGATGGAAATATCTGACGCTGGTGTGGCTCTCT 4630
QY ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysLys 1560
4631 GTTGGTTTGGCCCATACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4690
QY GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
4691 GGAGAGGATTCAGCCCGCTGAGTCTTTTGTGGCTCCATAAGCCAGCTCAACTCTCTGG 4750
QY AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
4751 GACTATGTCTGTCTCCACAGCAGGTTGAAGTCTAGCTACCTCTCTCCCGCAGAGAACTC 4810
QY SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys 1620
4811 AGTAAAGGAAACCTGTGTAGCATGGCTGATTTCTGTACAGAAATTTGTGGGAAAGTGAAG 4870
QY IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
4871 ATCGATTTAAAGACATATTTTGTCTGATTGCCCCACCGCTTAGGAGGCTCAGTGCCTCAT 4930
QY LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
4931 CTGAGAACTGCACTCTGAAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTCTGTGATCCA 4990
QY GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
4991 GGCTTTCAGCTGTGGGAAACCTGTGTCAGTACTGTCTGATTCAGGACAGCAGTGGACACA 5050
QY ProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPhe 1700
5051 CCATTTCTCCTCAGCTGAAACGCAATTAGCTGTGGGGTGCACCTCTCTTGGAGAAATGGCTTC 5110

Db	7271	CTCTGCCAACCTGGATGGACCTGGAGCTCTCCACTGCCAGAAATGTGTTCCAGTAGAATGT	7330
Qy	2441	ProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaIleIleLeuSer	2460
Db	7331	CCCCAACCTGAGGAATCCCAATGGAATCATTGATGTGCAAGGCCCTTGCCCTATCTCAGC	7390
Qy	2461	ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys	2480
Db	7391	ACAGCTCTCTATACCTGCAAGCAGGCTTTGAATTTGGTGGGAAATACATACCAACCCCTTGT	7450
Qy	2481	GlyGluAsnGlyHisTrpLeuGlyLysProThrCysLeuAlaIleGluCysLeuLys	2500
Db	7451	GGAGAAATGGTCACTGGCTCGAGGAAACCAACATGTAAAGCCATTGATGGCTTGAAA	7510
Qy	2501	ProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal	2520
Db	7511	CCCAAGGAGATTTGAAATGGCAATTTCTCTTACACGGACCTACACTATGGACAGACCGTT	7570
Qy	2521	ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu	2540
Db	7571	ACCTACTCTTGCACCGAGGCTTTGGCTCGAAGGTCCAGTGCCTTGACCTGTTTAGAG	7630
Qy	2541	ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln	2560
Db	7631	ACAGTGATTTGGATGTAGATGCCCATCTTGCATGGCCATCCACTGTGATTTCCCAAA	7690
Qy	2561	ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyr	2580
Db	7691	CCCATTTGAAATGGTTTTGTAGAGGTGCAGATTACAGCTATGGTGCCATATATCTAC	7750
Qy	2581	SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly	2600
Db	7751	AGTTGCTTCCCTGGGTTTCAGGTGGCTGTCATGCCATGCAGACCTGTGAAGTCAAGA	7810
Qy	2601	TrpSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp	2620
Db	7811	TGGTCAAGTTCATCCCAACATGTATGCCAATAGACTGTGGCTTCCCTCTCATATAGAT	7870
Qy	2621	PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspMet	2640
Db	7871	TTTGGAGACTGTACTAAACTCAAGATGACAGGGATATTTTGGACCAAGACGACATG	7930
Qy	2641	MetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThr	2660
Db	7931	ATGGAAATTTCCATATGTGACTCTCACCTCTCTTATCATTTGGGACAGTGCCTAAACC	7990
Qy	2661	TrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet	2680
Db	7991	TGGGAAATACAAAGGAGTCTCTGCTACACATTTCAATCAAACTTCTGTATGGTACCAG	8050
Qy	2681	ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGln	2700
Db	8051	GTTTTCATACACTGTATATCCAGGATATGAATCTTGGGGAACCTGTGTGCTATGCCAG	8110
Qy	2701	GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro	2720
Db	8111	GAAATGGAATCTGGAATGGCAGTGACCATCTCTGCAATTTCAATTTGAAATGTGACTTGCCT	8170
Qy	2721	ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln	2740
Db	8171	ACTGCTCTCGAAATGGCTTTTGGCTTTTACAGAGACTAGCATGGGAAGTGTCTGTGCGAG	8230
Qy	2741	TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn	2760
Db	8231	TATAGCTGTAAACCTGGACACATCTAGTGGGCTCTGACTTAAGGCTTTGTCTAGAGAT	8290
Qy	2761	ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro	2780
Db	8291	AGAAATGGAGTGGTGGCTTCCACGCTGTGAAGCCATTTCAATGCAAAAGCAATCCA	8350
Qy	2781	ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu	2800

Db	8351	GTTCATGAATGGATCCATCAAGGAACCAACTACACATACCTGAGCAGCTTGTACTATGAG	8410
Qy	2801	CysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn	2820
Db	8411	TGTGACCCCGATATGTGCTGAATGGCACTGAGAGGAGAACATGCCAGGATGACAAAAAC	8470
Qy	2821	TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla	2840
Db	8471	TGGGATGAGGATGAGGCCCATTTGCACTTCTGTGAGACTGCAAGTTCCACCCGCTCTGACC	8530
Qy	2841	AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn	2860
Db	8531	NATGGCCAGTGTGAGGAGACCGAGTACATTTCCAAAAGAGAGATTGAATACACTTGCAT	8590
Qy	2861	GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer	2880
Db	8591	GAAGGTTCTTCTTGGGGAGCCAGGAGTCCGGTTTGTCTTGGCCAATGGAAGTTGGAGT	8650
Qy	2881	GlyAlaThrProAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGly	2900
Db	8651	GGAGCCACTTCCCGACTGTGTGCTGTGCATGTGCCACCCCGCCACACACTGGGCATGGG	8710
Qy	2901	ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly	2920
Db	8711	GTGACGGAGGCTTGGACTATGGCTTTCATGAGGAAGTAACATTCCTCATGATGGGC	8770
Qy	2921	TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu	2940
Db	8771	TACATCTTTCAGCGTGTCTCCAAAACCTCACCTGTGCATGAGTGCACACTGGGATGCAGAG	8830
Qy	2941	IleProLeuCysLysProValAsnCysGlyProProGluAsnLeuAlaHisGlyPhePro	2960
Db	8831	ATTCTCTCTGTAAACCACTCACTGTGGACCTCTCTGAAGATCTTCCCAATGATTTCCCT	8890
Qy	2961	AsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyrLys	2980
Db	8891	NATGGTTTTCTTTATTCATGGGGCCATATACAGTATACAGTGTCTTCTCGTTATAAG	8950
Qy	2981	LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerPro	3000
Db	8951	CTCCATGGAAATTCATCAAGAAGTGCCTCTCCAAATGGCTCTCGAGTGGCACTCACCT	9010
Qy	3001	SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr	3020
Db	9011	TCTGCTGCTGCTGCAAGATGTTCCACCCAGTAATTAATGATATGGAATGTCTCATGGACA	9070
Qy	3021	AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly	3040
Db	9071	GATTTTGACTGTGGAAGGCGAGCCCGATTCAGTGTCTTCAAGGCTTCAAGCTCTTAGGA	9130
Qy	3041	LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu	3060
Db	9131	CTTTCGAATCACCTGTGAGCCGATGGCAGTGGAGCTCTGGGTTCGCCCACTGTGAA	9190
Qy	3061	HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer	3080
Db	9191	CACACTTCTTGTTGTTCTTTCATATGATACCAATGCGTTTCATCATGAGGACCACTCT	9250
Qy	3081	TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer	3100
Db	9251	TGGAGGAAAATGTGATACACTTACAGCTGCAGGTCTGGATATGTCATACAGGCGAGTTCA	9310
Qy	3101	AspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeu	3120
Db	9311	GATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCAGTGTGTGTAGCCCTTG	9370
Qy	3121	SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr	3140
Db	9371	TCTGTGGGTCCCAACCGTCTGTGCAATGCAATGCAATGCAATGCAATGCAATGCAAT	9430
Qy	3141	GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThr	3160
Db	9431	GAAAGTAGAGTGAATCTCAGATGTCTGGAAGGTTATACGATGATACAGATACAGATACAG	9490

QY 3161 PheThrCysGlnLysAspGlyArgTTPheProGluArgIleSerCysSerProLysLys 3180
 DB 9491 TTCACCTGTCAGAAAGATGCTGCTGCTCCCTGAGAGAAATCTCTGCACTCTCAAAAAA 9550
 QY 3181 CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
 DB 9551 TGTCCTCTCCCGGNAACATAACACATATACTTGTTCATGCGGACGANTTTCAGTGTGAAT 9610
 QY 3201 ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
 DB 9611 AGCAAGATTCTGTGTCATGTCAGAGGGGTATACCTTTTCAGGCGAGTTAATACATACAGTA 9670
 QY 3221 CysGlnLeuAspGlyThrTTPGluProProPheSerAspGluSerCysSerProValSer 3240
 DB 9671 TGTCACTTGATGGAACCTGGAGGACCAACCAATCTCCGATGAATCTTGCACTCCAGTTTCT 9730
 QY 3241 CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu 3260
 DB 9731 TGTGGNAACCTGAAAGTCCAGAACATGGATTTGTGGTTCGCAGTAATACACCTTTGAA 9790
 QY 3261 SerThrIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgVal 3280
 DB 9791 AGCACAATATTATCAGTGTGAGCTGCTGCTATCAACTAGAGGGGAACAGGCAACGCGTC 9850
 QY 3281 CysGlnGluAsnArgGlnTTPSerGlyGlyValAlaIleCysLysGluThrArgCysGlu 3300
 DB 9851 TGCAGAGAGAACAGACAGTGTGAGTGGAGGGGTGCAATATGCAAGAGACACAGTGTGAA 9910
 QY 3301 ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrGlyProAsn 3320
 DB 9911 ACTCCACTTGAATTTCTCAATGGGAAGCTGACATTTGAAACAGGACGACTGGACCCAC 9970
 QY 3321 ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr 3340
 DB 9971 GTGGTATATTCTGTCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCACACTGCACA 10030
 QY 3341 GluAsnGlyThrTTPSerHisProValProLeuCysLysProAsnProCysProValPro 3360
 DB 10031 GAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGCAAAACAAATCATGCCCTGTTCCT 10090
 QY 3361 PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal 3380
 DB 10091 TTTGTGATTTCCCGAATGCTCTGCTCTGAAAGGGAGTTTTATGTTCATCAGAAATGTG 10150
 QY 3381 SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyLleIleThrCysAsnPro 3400
 DB 10151 TCCATCAAAATGTAGGGAAGGTTTTCTGTGCGAGGGCCACGGCATCATTTACCTGCAACCCC 10210
 QY 3401 AspGluThrTTPThrGlnThrSerAlaLysCysGluLysLysIleSerCysGlyProProAla 3420
 DB 10211 GACGAGCGTGGACACAGACAGCGCCAAATGTGNAANAATCTCAATGTGTCCACAGCT 10270
 QY 3421 HisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
 DB 10271 CACGTAGAAATGCAATTTGCTCGAGGGGTACATTTATCAATATGAGACATGATCACCTAC 10330
 QY 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly 3460
 DB 10331 TCATGTTTACAGTGGATACATGTTGTGAGGGGTTTCTGTGAGGAGTGTTCCTGTAGAAATGGA 10390
 QY 3461 ThrTTPThrSerProProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGly 3480
 DB 10391 ACATGGACATCACTCTCTATTTGCGAGAGCTGTCTGTGATTTCAATGTCAGATGGGGC 10450
 QY 3481 IleCysGlnArgProAsnAlaCysSerCysProGluGlyTTPMetGlyArgLeuCysGlu 3500
 DB 10451 ATCTGCCAACGCCCAATGCTTGTCTGTCTGTCAGAGGGCTGGATGGGGGCTCTGTGAA 10510
 QY 3501 GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys 3520
 DB 10511 GAACCAATCTGCATTTCTTCCCTGTCTGTGACAGGAGTCTGCTGTGTGGGCCCTTACCACTGT 10570

QY 3521 AspCysProGlyTTPThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
 DB 10571 GACTGCCCGCTGCTGGACGGGCTCTCGCTGTATACAGCTGTTTCCAGTCTCCCTGC 10630
 QY 3541 LeuAsnGlyGlyLysValArgProAsnArgCysHisCysLeuSerSerTTPThrGly 3560
 DB 10631 TTAATAGTGTGAAATGCTGTAAAGACCAACCGATGCTCACTGTCTTCTTCTTGGACGGGA 10690
 QY 3561 HisAsnCysSerArgLysArgArgThrGlyPhe 3571
 DB 10691 CATAACTGTTCCAGGAAAGAGGACTGGGTTT 10723
 RESULT 2
 AAD57253
 ID AAD57253 standard; cDNA; 11546 BP.
 XX
 AC AAD57253;
 XX
 DT 06-NOV-2003 (first entry)
 XX
 DE Human CGDD-33 cDNA.
 XX
 KW Human; cell growth, differentiation and death protein; CGDD; leukaemia;
 KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
 KW muscular disorder; myotonic dystrophy; catatonia; endocrine disorder;
 KW diabetes; Grave's disease; cancer; immunological disorder; scleroderma;
 KW systemic lupus erythematosus; allergy; Crohn's disease; renal disorder;
 KW gastrointestinal disorder; Goodpasture's syndrome; infection; cirrhosis;
 KW cardiovascular disorder; atherosclerosis; hepatic disease; transgenic;
 KW transgenic animal; gene therapy; neuroprotective; relaxant; cytostatic;
 KW dermatological; immunosuppressive; cerebroprotective; anticonvulsant;
 KW antibacterial; antiparasitic; fungicide; virucide; uropathic; cardiant;
 KW protozoacide; nootropic; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 298..11010
 FT /*tag= a
 FT /product= "Human CGDD-33 protein"
 XX
 PN WO2003050253-A2.
 XX
 PD 19-JUN-2003.
 XX
 PF 04-DEC-2002; 2002WO-US039133.
 XX
 PR 07-DEC-2001; 2001US-0340747P.
 PR 20-DEC-2001; 2001US-0342761P.
 PR 15-JAN-2002; 2002US-0349705P.
 PR 06-FEB-2002; 2002US-0354764P.
 PR 12-FEB-2002; 2002US-0356216P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Griffin JA, Ramkumar J, Emerling BM, Kable AE, Elliott VS,
 PI Marquis JP, Baughn MR, Gorvad AE, Yue H, Lee BA, Becha SD, Tang YT;
 PI Tran UK, Swarnakar A, Lee S, Ison CH, Hafalia AJA, Tran B;
 PI Sprague WW, Lee SY, Khare R, Gandhi AR, Gietzen KJ, Bhatia U;
 PI Burhill JD, Blake JJ, Ho A, Zheng W;
 XX
 DR WPI; 2003-532903/50.
 DR P-PSDB; AAD37944.
 XX
 FT New CGDD polypeptides, useful for diagnosing, preventing, and treating
 FT disorders associated with an abnormal expression or activity of CGDD,
 FT e.g. neuromuscular, immunological, cardiovascular disorders, cancer
 FT and/or infections.
 XX
 PS Claim 5; Page 296-299; 299pp; English.
 XX
 CC The present invention relates to novel cell growth, differentiation and
 CC death (CGDD) proteins and polynucleotides encoding them. The sequences of

CC the invention are useful in diagnosing, preventing and treating disorders
 CC associated with an abnormal expression or activity of CGDD such as
 CC neurodegenerative disorders (e.g. Parkinson's disease, Alzheimer's
 CC disease), muscular disorders (e.g. myotonic dystrophy, catatonial),
 CC endocrine disorders (e.g. diabetes, Grave's disease), cancers (e.g.
 CC leukaemia, cervical or breast cancers), immunological disorders (e.g.
 CC scleroderma, systemic lupus erythematosus, allergies), gastrointestinal
 CC disorders (e.g. Crohn's disease), renal disorders (e.g. Goodpasture's
 CC syndrome), infections (e.g. viral, bacterial, fungal, parasitic,
 CC protozoal, helminthic), cardiovascular disorders (e.g. atherosclerosis)
 CC and hepatic diseases (e.g. cirrhosis). The polynucleotides can be used to
 CC create humanised animals or transgenic animals to model human diseases.
 CC The invention is also used in gene therapy. The present sequence is human
 CC CGDD-33 cDNA

XX Sequence 11546 BP; 3008 A; 2726 C; 2836 G; 2976 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 11546
 Score: 19905.50 Matches: 3561
 Percent Similarity: 99.80% Conservative: 3
 Best Local Similarity: 99.72% Mismatches: 6
 Query Match: 99.66% Indels: 1
 DB: 8 Gaps: 1

US-09-977-053-4 (1-3571) x AAD57253 (1-11546)

QY 1 MetTProArgLeuAlaPheCysTyrGlyLeuAlaLeuValSerGlyTProAlaThr 20
 DB 298 ATGTGGCCCTCGCCCTGCGCCCTTTTGTCTCGGGGTCTGGCGCTGTTTCGGGCTGGGGCACC 357
 QY 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
 DB 358 TTTTCAGCAGATGTCCTCGCTCGCCATTTTCAGCTTCGGCTCTTCCCGCAGACCGCGGCC 417
 QY 41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
 DB 418 GGGGGCCCCGGGAGTATCCCGCGCCCGCTCTCTGGCGAGCGAGCGCGGGAGCAGA 477
 QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
 DB 478 GTGGAGCGGCTGGGCGAGCGGTTCGGCGAGCGGTCTGGCGAGCTCAGCGAG 537
 QY 81 ArgLeuGluLeuValPheLeuValAspSerSerValGlyGluValAlaAsnPheArg 100
 DB 538 CGCTTGGAGCTTGTCTTCTGTGTGATGATTCGTCCAGCGTGGCGAAGTCAACTTCGCG 597
 QY 101 SerGluLeuMetPheValArgIlysLeuLeuSerAspPheProValValProThrAlaThr 120
 DB 598 AGCGAGCTCATGTTCTCGCGAAGCTGCTCTCGACTTCCCGTGTGGTCCCGCGCGAG 657
 QY 121 ArgValAlaIleValThrPheSerSerIysAsnTyrValValProArgValAspTyrIle 140
 DB 658 CGCGTGGCCATCGTGACCTCTCTGTCCAGAACTACGTGTGGTGGCGCGTCTGATACATC 717
 QY 141 SerThrArgAlaArgGlnHisIlysCysAlaLeuLeuGlnGluIleProAlaIle 160
 DB 718 TCCACCGCGCGCGCGCGCAGCAGTGGCGCTCTCTCCAGAGATCCCTGCCATC 777
 QY 161 SerTyrArgGlyGlyThrThrIlysGlyAlaPheGlnGlnAlaAlaGlnIleLeu 180
 DB 778 TCTACCGAGGTGGCGGCACCTACACCAAGGGCGCTTCAGCAAGCGCGCGCAAAATCTT 837
 QY 181 LeuHisAlaArgGluAsnSerThrIlysValIlePheLeuIleThrAspGlyTyrSerAsn 200
 DB 838 CTTTCATCGTAGAGAAACTCAACAAAGTTGTATTTCTCATCTACTGTGATATTCAT 897
 QY 201 GlyIAspProArgProIleAlaIleSerLeuArgAspSerGlyValGluIlePheThr 220
 DB 898 GGGGAGACCCCTAGACCAATTCGAGGTCTACCTCGAGATTCAGAGTGGAGATCTTCACT 957
 QY 221 PheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIysGlu 240

DB 958 TTTGGCATATGCAAGGGAAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 1017
 QY 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260
 DB 1018 GAGCAGCTGTTACTGCTACACAGTTTGAAGAATTTAGGCTTTAGCTCGCGGCAATG 1077
 QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
 DB 1078 CATGAAGATCTACCTTCGGGAGTTTATTTCAAGATGATATGCTCCACTGCTCATATCTT 1137
 QY 281 CysAspGluGlyIysAspCysCysAspArgMetGlySerCysIysCysGlyThrHisThr 300
 DB 1138 TGTGATGAAGGCGAGAGACTGCTGTGACCAGAAATGGGAAGCTGCAAAATGTGGGACACACA 1197
 QY 301 GlyHisPheGluCysIleCysGluIysGlyTyrTyrGlyIysGlyLeuGlnTyrGluCys 320
 DB 1198 GGCCATTTTGAATGTCATCTGTGNAAGGGGTATTACGGGNAAGTCTGCAGTATGAATGC 1257
 QY 321 ThrAlaCysProSerSerGlyThrTyrIlysProGluGlySerProGlyIlysSerSerCys 340
 DB 1258 ACAGCTTGGCCATCGGGACATACAAACCTGAAGGCTCACCAGGAGGAATCAGCAGTTGC 1317
 QY 341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
 DB 1318 ATTCCATGTCCTGATGAATAATCACCTCTCCACCTGGAGACATCCCTCTGAAGACTGT 1377
 QY 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
 DB 1378 GTCTGCAGAGGGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGCCACTGCGCTGCC 1437
 QY 381 LeuIlysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
 DB 1438 CTGAAGCCTCCCGAAATGGTTACTTTATTCAAACACTTGCACCAACCACTTCAATGCA 1497
 QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
 DB 1498 GCTGTGGGGTCCGATGTCACTCTGATTTGATCTTGTGGAGAGCAGCATCATCTTATGT 1557
 QY 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
 DB 1558 CTACCCAAATGGTTGTGGTCCGCTTCAGAGAGCTACTGCAGAGTAAGAACATGTCTCAT 1617
 QY 441 LeuArgGlnProIysHisIleSerCysSerThrArgGluMetLeuTyrIysThr 460
 DB 1618 CTCCGCGACCGCAACATGCGCACATCTGTCTTCAAGGGAAATGTTATATAAGCA 1677
 QY 461 ThrCysLeuValAlaCysAspGlyTyrArgLeuGluGlySerAspIysLeuThrCys 480
 DB 1678 ACATGTTTGTGCTTGTGTATGAGGGTACAGACTAGAGGCGAGTATAGCTTACTTGT 1737
 QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
 DB 1738 CAAGGAAACAGCCAGTGGCATGGCCAGAACCCCGGTGTGTGGAGCGCCACTGTTCACCC 1797
 QY 501 PheGlnMetProIysAspValIleIleSerProHisAsnCysGlyIysGlnProAlaIys 520
 DB 1798 TTTTCAGATCCCAAGATGTTCATCATATCCCCCACAACCTGTGGCAGCAGCAGCCAAA 1857
 QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValIysGlu 540
 DB 1858 TTTGGACGATCTGCTATGTAGTTCGCCCAAGGGTTCATTTTATCTGGAGTCAAGAA 1917
 QY 541 MetLeuArgCysThrThrSerGlyIysTrpAsnValGlyValGlnAlaIleValCysIys 560
 DB 1918 ATGCTGAGATGTACCACTTCGGAATAATGGAATGTGGAGTTCAGGCGAGCTGTGTGAAA 1977
 QY 561 AspValGluAlaProGlnIleAsnCysProIysAspIleGluAlaIysThrLeuGluGln 580
 DB 1978 GAGTGGAGGCTCTCAAAATCAACTGTCTTCAAGACATAGAGCTTAGACTCTTGGACAG 2037
 QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaIysAspAsnSerGlyIysGlu 600
 DB 2038 CAAGATTCCTGCCAATGTTTACCTGGCAGATTCCAAACAGCTAAAGACACTCTGTGGTAAAAG 2097

601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
 2098 GTGTCAGTCACAGGTTTCATCAGCTTTTCACCCACCTTACCTTTTCCCAATGGAGATGT 2157
 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
 2158 GCTATCGTATACAGCGCACTGACCTATCCGGCAACCAAGCCAGCTGCATTTTCCATATC 2217
 641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
 2218 AAGGTTATTGATGACAGAACCACTGTCATAGACTGGTGCGAGATCCACCTCCCGTCCAG 2277
 661 ValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
 2278 GTCTCGGAGAGGTACATGCGCGAGCTGGAGTGAAGCTCAGTTCTCAGACAACTCAGGG 2337
 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
 2338 GCTGAATGGTCATTACAGAGTATACACAGGAGACCTTTTCCCTCAAGGGGAGACT 2397
 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
 2398 ATAGTACAGTATACAGCACTGACCCCTCAGGCCAATAACAGGACATGTGATATCCATATT 2457
 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
 2458 GTCATAAAGGTTCTCCCTGTGAATATCCATTACACCTGTAAATGGGATTTTATATGC 2517
 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
 2518 ACTCCAGATAACTCGAGTCAACTGTATATTAATTCCTTGGAGGGCTATGATTTCACA 2577
 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
 2578 GAAGGGTCTACTGACAAAGTATTATGTGCTTATGAAGATGGCGTCTGGAAACCAACATAT 2637
 781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
 2638 ACCACTGAATGGCGAGACTGTGCCAANAACGTTTTCANAACCAAGGGTTCAAGTCTCTTT 2697
 801 GluMetPheTyrLysAlaLysArgCysAspThrAspLeuMetLysLysPheSerGlu 820
 2698 GAGATGTTCTACAAAGCAGCTCGTTGTGTATGATGACACAGATCTGAAGAAGTTTCTGAA 2757
 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
 2758 GCATTTTGAGACGACCTCGGGAAAAATGGTCCCATCATTTTGTATGATGACAGAGCAATT 2817
 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
 2818 GACTGCAAGACTGGAGGAGAACCTGACCAANAATTTTGGCTAGATATATATTATGACTAT 2877
 861 GluAsnGlyPheAlaIleGlyProGlyTyrProGlyAlaAlaAsnArgLeuAspTyrSer 880
 2878 GAAATGGCTTTGCAATTTGACACAGGTGGCTGGGGTGCAGCTAATAGCTGGATTACTCT 2937
 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
 2938 TACGATGACTTCTCGGACACTGTGCAAGAAACAGCCCAAGGATCGGCAATGCCAAGTCC 2997
 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
 2998 TCACGGATTAAGAAGTGGCCCAATTAATCTGACTATTAANAATTAAATTTTAAATC 3057
 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGln 940
 3058 ACAGTAGTGTGCAATTTACCGATGAAGAAATGATACCTTGAATGGGAAAAATCAGCAA 3117
 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
 3118 CGACTCCTTCAGACATTTGGAAACTATCACAAATAAACTGAAAGGAGCTCTCAACAAGAC 3177

961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
 3178 CCCATGATTCTCTTTCAGCTTGCATCAGAAATCTTATAGCCGACAGCAATTCATTAGAA 3237
 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
 3238 ACAAAAAGGCTTCCCTCTCTGAGACACAGGCTCAGTGTCTGAGAGGGCGTATGTGTGTC 3297
 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
 3298 AATTGCCCTTTGGGAACCTATTAACTGGAACATTTTCACTGTGAAGCTGCCGGATC 3357
 1021 GlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
 3358 GGATCTTATCAAGATGAAGAAGGGCACTTGAGTGAAGCTTTGCCCTCTGGGATGATC 3417
 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
 3418 ACGAATATATTCATTCAAGAAACATCTCTGATTGTAAGCTCAGTGTAAAACAGGCACC 3477
 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
 3478 TACTCATACAGTGGACTTGAAGTGTGCTCCACTGGGCACCTTATCAGCCAAA 3537
 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
 3538 TTTGGTTCCCGAGCTGCTCTCTGTGTCAGAAAACACTCACTGTGAAAAGAGAGGCC 3597
 1101 ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu 1120
 3598 GTGAACATTTCTGCATGTGGAGTTCCTGTGTCAGAAAGGAAAATTCCTCGGTCTCGGTTA 3657
 1121 MetProCysHisProCysProAspArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys 1140
 3658 ATGCCCTGTACCCATGCTCTCGTACTATTAACCAACCTAATGAGGAGAGGCCCTTCGCG 3717
 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys 1160
 3718 CTGGCTGTCTCTTTTATGGAACCTACCCCATTCGCTGGTTCAGATCGCTCACAAATGT 3777
 1161 SerSerPheSerSerThrPheSerAlaAlaGluLysValValProProAlaSerLeu 1180
 3778 TCAAGTTTATAGTTCACTTTCTACGGCGCAGAGAAAGTGTGGTGCCTCTCTCTCTCT 3837
 1181 GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn 1200
 3838 GGCATATTAAAGAGCGATGAATTCAGCACTCAGGTTTTCCATGAATGCTTCTTTAAC 3897
 1201 ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro 1220
 3898 CTTGGCCACAATAGTGGAACTGCGCAGCAACTTGGGGGTGGTATTGTTGTTCTCTGTCCA 3957
 1221 LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys 1240
 3958 CTTGGATATACAGGCTTAAAGTGTGAACAGACATCGATGATGTCAGCCCACTGCTTTCG 4017
 1241 LeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGly 1260
 4018 CTCACAAATGGAGTTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTGGCCATCAGGT 4077
 1261 TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn 1280
 4078 TACACAGGTGAGCGTGTGAAGAAAATATAATCAGTGTAGCTCCAGTCTCTGTTTAAAT 4137
 1281 LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal 1300
 4138 AAAGGAATCTGTGTGTGTGGTGTGGCTATGCTGTCACATGTGTGAAAAGGATTTGTA 4197
 1301 GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla 1320
 4198 GGCCTGATGTTGAACAGAGTCAATGAATGCCAGTCAAAACCCATGCTTTAATTAATGCA 4257
 1321 ValCysGluAspGlnValGlyPheLeuCysLysCysProProGlyPheLeuGlyThr 1340

Db 4258 GTCTGTGAAGACAGAGTTGGGGATTTCTTGTGCAATATGCCACCTGGATTTTGGGTACC 4317
QY 1341 ArgCysGlyLeuValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys 1360
Db 4318 CGATGTGGAAAGAACGTGATGAGTGTCTCAGTCAGCCATGCAAAATGGAGCTACCTGT 4377
QY 1361 LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
Db 4378 AAAGACGGTGCATAGCTTCAGATGCCCTGTGTGCAGCTGGCTTCACAGATCACACTGT 4437
QY 1381 GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
Db 4438 GAATTTGACATCAATGAATGTAGTCTAATCAATGATAGAAATCAGCCCACTGTGTGGAT 4497
QY 1401 GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr 1420
Db 4498 GAATTAATTCATACAGTTGTAAATGTCAGCCAGGATTTTCAGGCCAAAGAGTGTGAACA 4557
QY 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
Db 4558 GAACAGTCTACAGGCTTTAAACCTGGATTTTGAAGTTTCTGGCATCTATGGATATGTGATG 4617
QY 1441 LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSer 1460
Db 4618 CTAGATGGCATGCTCCCATCTCTCCATGCTCTTAACCTGTACCTCTGTGGATGAATCTCT 4677
QY 1461 AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr 1480
Db 4678 GACGACATCAACTATGGAACACCAATCTCTATGCGAGTTGTATAAACGCGACGACATACC 4737
QY 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThr 1500
Db 4738 TTGCTCTCTGACTGATTAATACCGCTGGGTCTTTATGTGAATGGCGAGGAAGATAACA 4797
QY 1501 AsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrSerAla 1520
Db 4798 AACTGTCCCTCGTGAAATGATGGCAGATGGCATCATATTGGCATCACTTGGACAGTGCC 4857
QY 1521 AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer 1540
Db 4858 AATGGCATCTGGAAGTCTATATCGATGGGAAATATATCTACGGTGGTCTGGCTCTCT 4917
QY 1541 ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLysIys 1560
Db 4918 GTTGGTTTCCCATACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4977
QY 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyr 1580
Db 4978 GGAGAGGATTCAGCCAGCTGATCTTTTGTGGGCTCCNTAAGCCAGCTCAACCTCTGG 5037
QY 1581 AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeu 1600
Db 5038 GACTATGCTCTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTCTGCCAGAGGAATC 5097
QY 1601 SerLysGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyValValLys 1620
Db 5098 AGTAAAGGAACGTTGTAGCATGGCTGATTTCTGTCAAGAAATGTGGGAAAGTGAAG 5157
QY 1621 IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis 1640
Db 5158 ATCGATTCATAGAGCATATTTGTCTGATGCTCCACGCTTAGGAGGGTCAGTGCCTCAT 5217
QY 1641 LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro 1660
Db 5218 CTGAGAACTGCATCTCAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTCTGTGATCCA 5277
QY 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
Db 5278 GGCCTTCAGCTGGTGGGAACCTGTGCGAGTACTGTCTGAATCAAGGACAGTGGACAAA 5337
QY 1681 ProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPhe 1700

Db 5338 CCACCTCTCTCCTACCTGTGAACGCATTAGCTGTGGGGTGCACCTCTCTTTGGAGAAATGGCTTC 5397
QY 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
Db 5398 CATTCAGCGATGACTTCTATGTGTCAGCACAGTAACCTTACCAAGTGCACAACTATGGCTAC 5457
QY 1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSer 1740
Db 5458 TATCTATTTGGTGTACTCAAGGATGTTCTGTACAGATTAATGGGAGCTTGGAAACGGGTTTCA 5517
QY 1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
Db 5518 CCATCTCTCTCTGATGTGATGATGTCAGTTGGATTCAGATTTGTAGTGAGCATGCTTCT 5577
QY 1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
Db 5578 TGCCTGAAACGTAGATGATCTTACATATGTTCTATGTGTCCACCGTACACAGAGATGGG 5637
QY 1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5638 AAAAATCTGTGCAGAACCTATTAATGTAAAGCTCCAGGAAATCCGGAAATATGGCCACTCC 5697
QY 1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
Db 5698 TCAGGTGAGATTTATACAGTAGTGCCGCGAGTCACATTTTGTGTGAGGAAGATACACAG 5757
QY 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluThrPheAsnHisLeuIlePro 1840
Db 5758 TTGATGGGAGTAAACCAAAATCACTGTTGGAGTCTGGAGAAATGGAATTCATCTAATACCA 5817
QY 1841 TyrCysLeuAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
Db 5818 TATTTGAAGCTGTTTCACTGTTGGTAAACCGGCTATTTCCAGAAATATGTTCCATTGAGGAG 5877
QY 1861 LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla 1880
Db 5878 TTAGCATTTTACTTTTGGCAGCAAGTGCATATAGTGTAAATAAGGATATCTCTGGCC 5937
QY 1881 GlyAspLysGluSerSerCysLeuAlaAsnSerSerTyrSerHisSerProValCys 1900
Db 5938 GGTGATAAAGATCATCTCTGTCTTGTAAACAGTCTTGTGAGTCAATCTCCCTCTCTGTGT 5997
QY 1901 GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly 1920
Db 5998 GAAACGTCAGTGTCTAGTCCGGAATATTAATAATGGAATAATATATTTTGGTGGG 6057
QY 1921 LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro 1940
Db 6058 CTTACCTACTCTTCTACTGTCATCATATTCTGCGATACAGGATACAGCTTACAGGGCCCT 6117
QY 1941 SerIleIleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHisLeu 1960
Db 6118 TCCATTTATTAATGACCGGCTTCTGGCATCTGGGACAGAGCGCCACTCTGCTGTACCTC 6177
QY 1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
Db 6178 GTCTTCTGTGGAGAACCACTGTCATCAAGATGCTGTCTATTACGGGGAATAAATCACT 6237
QY 1981 PheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr 2000
Db 6238 TTCAGGAACACCGCTCACTTACACCTTGCATAAGAGGCTATATCTTGTGTGTCTTTCACAC 6297
QY 2001 IleGluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaValSer 2020
Db 6298 ATTGAATGCTGGCCGACCGGCAAGTGGAGTAGAGTAGACAGCAGTGCCTGGCTGTCTCC 6357
QY 2021 CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly 2040
Db 6358 TGTGATGACCACTTCTGACCACTCTTCCAGAGACTGCCCATCGGCTCTTTTGA 6417
QY 2041 AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu 2060
Db 6418 GACATTTGCACTTCTACTACTGCTCTGATGGTTACAGCTTAGCAGCAATATCCCGACTTCTC

QY	2061	CysAenAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAla	2080
DB	6478	TGCAATGCCCAAGGGCAAGTGGGTACCCAGAGAGGTCAAGACATGCCCCGTGTGTAGCT	6537
QY	2081	HisPheCysGluLysProProSerValSerTrpSerIleLeuGluSerValSerLysAla	2100
DB	6538	CAATTCTGTGAAAAACCTCCATCGCTTCTCTATAGCATCTTGGAACTCTGTGAGCAAGCA	6597
QY	2101	LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr	2120
DB	6598	AAATTTGACAGCTGGCTCAGTTGTGAGCTTTAAATGCATGGNAGGCTTGTACTGAACACC	6657
QY	2121	SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln	2140
DB	6658	TCAGCAAGATTTGAATGTATGAGAGGTGGGCAGGTGGAAACCTTCCCCCATGTCCATCCAG	6717
QY	2141	CysIleProValArgCysGlyGluProProSerIleMetAenglyTrpAlaSerGlySer	2160
DB	6718	TGCATCCCTGTGCGGTGTGGAGAGCCACCAAGGCATCATGAATGGCTATGCAAGTGGATCA	6777
QY	2161	AsnTrpSerPheGlyAlaMetValAlaTrpSerCysAsnLysGlyPheTrpIleLysGly	2180
DB	6778	AACTACAGTTTGGAGCCATGGTGGCTTACAGCTGCAACCAAGGGGTTCTACATCAAGGG	6837
QY	2181	GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis	2200
DB	6838	GAAGAAGAGAGCACCTGCGCAAGCCACAGGGCAGGTGGAGTAGTCCCTATACCGAGTGGCCAC	6897
QY	2201	ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly	2220
DB	6898	CCGGTATCTTGTGTGTGAACACCTTAAAGGTTGAGAAATGGCTTCTCGAGACATACAACTGGC	6957
QY	2221	ArgIlePheGluSerGluValArgTrpGlnCysAsnProGlyTrpLysSerValGlySer	2240
DB	6958	AGGATCTTTGAGAGTGAAGTGAGGTATCAGTGTAAACCCGGGCTATATAAGTCAGTCGGAAGT	7017
QY	2241	ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal	2260
DB	7018	CCCTGTATTTGTCTGCCAAGCCAAATCGCCACCTGGGCACAGTAGTAATCCCTCTCGATGTGTGT	7077
QY	2261	ProLeuAspCysGlyLysProProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe	2280
DB	7078	CCCTCGACGTGTGAAACCTTCCCCCGATCCAGAAATGGCTTATGAAGAAGAGAAACTTT	7137
QY	2281	GluValGlySerLysValGlnPhePheCysAsnGluGlyTrpGluLeuValGlyAspSer	2300
DB	7138	GAAGTAGGGTCCAAAGTTTCACTTTTCTGTAATGAGGGTTATGAGCTTGTGTGTGACAGT	7197
QY	2301	SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro	2320
DB	7198	TCTTTGGACATGTCAGAAATCTGGCAAAATGGAATTAAGAAGTCAAAATCCAAAGTGCATGCCCT	7257
QY	2321	AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThr	2340
DB	7258	GCCAAAGTGCCAGAGCCGCCCTCTTGGAAAAACCAAGCTAGTATTTAAAGGAGTGTGACCAACC	7317
QY	2341	GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal	2360
DB	7318	GAGGTAGGAGTTGTGACATTTCTGTAAAGAAGGGCATGCTCTGCAAGGCCCTCTGTGC	7377
QY	2361	LeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleValLeu	2380
DB	7378	CTGAATATGCTTGGCCATCCCAAGCAATGGAATGACCTTTTCTTCCCTGTTGTAGAATGTGTCTT	7437
QY	2381	CysThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe	2400
DB	7438	TGTAACCCCAACTCCCTTAATTTCTTGTGTGTGCCATTCCTTCTTCTCTCTCTCTCTCTCT	7497
QY	2401	GlySerThrValLysTrpSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThr	2420
DB	7498	GGAACTACTGTCAAGTATCTTGTGTAGGTGGGTTTTTCTTAAGAGAAATTTCTACCAACC	7557

Qy	2421	LeuCyGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys	2440
Db	7558	CTCTGCCAACCTGATGGCACCTGGAGCTCTCCACTCCCAAGATGTGTTCAGTAGAATGT	7617
Qy	2441	ProGlnProGluLeuProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSer	2460
Db	7618	CCCCACCTGAGGAAT---CCCCATGAATCAATTTGATGTGAGGCCCTTGCTATCTCAGC	7674
Qy	2461	ThrAlaLeuTyrThrCysIysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys	2480
Db	7675	ACAGCTCTCTATACCTGCAGCCAGGCTTTGAATTTGGTGGGAATACTACCCACCTTTGT	7734
Qy	2481	GlyGluAsnGlyHisTrpLeuGlyGlyIysProThrCysIysAlaIleGluCysLeuIys	2500
Db	7735	GGAGAAAATGGTCACTGGCTTCGAGGAAAACCAACATGTAAGGCCATTCAGTGCCTGAAA	7794
Qy	2501	ProIysGluIleLeuAsnGlyIysPheSerTyrThrAspLeuHisTyrGlyGlnThrVal	2520
Db	7795	CCCAAGAGAAATTTGAATGGCAAAATTCCTTTACAGGACCTACACTATATGACAGACCGTT	7854
Qy	2521	ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu	2540
Db	7855	ACCTACTCTTGCACCGAGGCTTTCGGCTCGAAGGTCCCACTGCCTTGACCTGTTAGAG	7914
Qy	2541	ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln	2560
Db	7915	ACAGGTGATTTGGGATGTAGATGCCCATCTTGCATGCGCATCCACTGATTCGCCACAA	7974
Qy	2561	ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyr	2580
Db	7975	CCCATTTGAAAATGGTTTTTGTAGAAGGTGCAGATTACAGCTATAGGTGCCATAATCTCTAC	8034
Qy	2581	SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly	2600
Db	8035	AGTTGCTTCCTCGGCTTCAGGTGGCTGATCATGCCACAGACCTGTGAGAGTCCAGGA	8094
Qy	2601	TrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp	2620
Db	8095	TGGTCAAGTTTCCATCCCAACATGTATGCCAATAGACTGTGGCCCTCCCTCATATAGAT	8154
Qy	2621	PheGlyAspCysThrIysLeuIysAspAspGlnGlyTyrPheGluGlnGluAspAspMet	2640
Db	8155	TTTGGAGACTGTACTAAACTCAAAGATGACACGGGATATTTTGACGACGAGACGACATG	8214
Qy	2641	MetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaIysThr	2660
Db	8215	ATGGAGTTTCATATGTGACTCTCTACCCCTCCTTATCATTTGGGACGAGTGGCTAAACC	8274
Qy	2661	TrpGluAsnThrIysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMet	2680
Db	8275	TGGGAAAATACAAAGGAGTCTCTGCTACACATTCATCAAACTTTCTGTATGTGATCAGTG	8334
Qy	2681	ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValleuIleCysGln	2700
Db	8335	GTTTTCAACACCTGTAAATCCAGATATGAACCTTCGGGAAACCCCTGTGCTGATCTGCCAG	8394
Qy	2701	GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro	2720
Db	8395	GAGATGGAACTTTGGAAATGGCAGTGCACCATCTCGCATTTCAATTGAAATGTGACTTGCCT	8454
Qy	2721	ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln	2740
Db	8455	ACTGCTCTCGAAAATGGCTTTTTCGGTTTTTACAGAGACTAGCATGGGAAGTGTGTGCGAG	8514
Qy	2741	TyrSerCysIysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn	2760
Db	8515	TATAGCTGTAAACCTGGACACATTCCTAGCAGGCTCTGACTTAAGGGCTTTGTCTAGAGAA	8574
Qy	2761	ArgGlyTrpSerGlyValSerProArgCysGluAlaIleSerCysIysValysProAsnPro	2780
Db	8575	AGAAATGGAGTGGTGCCTCCCCACGCTGTGAAGGCCATTTTCATGCGAAAAGCCAAATCCA	8634
Qy	2781	ValMetAsnGlySerIleIysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu	2800

Db	8635	GTCATGAATGGATCCCATCAAGAGAAAGCAACTACACATACCTGAGCAGCTGTACTATGAG	8694
Qy	2801	CysAspProGlyTyrValLeuAsnGlyThrGluArgThrCysGlnAspAspLysAsn	2820
Db	8695	TGTGACCCCGGATATGTGCTGAATGCACTGAGAGGAGAACATGCCAGGATGACAAAC	8754
Qy	2821	TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla	2840
Db	8755	TGGGATGAGGATGAGCGCCATTTGCAATTCCTGTGACTGCGAGTTCACTCCCGGCTCAGCC	8814
Qy	2841	AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn	2860
Db	8815	AATGGCCAGGTGAGGAGACGAGTACATTCCAAAGAGATGATACACTTCGAAT	8874
Qy	2861	GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer	2880
Db	8875	GAAGGGTCTCTGCTTGAGGGAGCCAGAGTGGGTTTGTCTTGCCAAATGGAGTTGGAGT	8934
Qy	2881	GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly	2900
Db	8935	GGAGCCACTCCCGACTGTGTGCTGTGCATGTGCACCCCGCCACAACCTGGCCAAATGGG	8994
Qy	2901	ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly	2920
Db	8995	GTGACGGAAGGCTTGACTATGGCTTCATGAAGAGAGTAACTTCCACTGTACAGAGGC	9054
Qy	2921	TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu	2940
Db	9055	TACATCTTCGCGGTCTCCAAACTCACCTGTCTGCTGAGTGGCAACTGGGATGCAGAG	9114
Qy	2941	IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro	2960
Db	9115	ATTCTCTCTGTAAACAGTCAACTGTGACCTCTCGAAGATCTTGCCCATGTGTTCCCT	9174
Qy	2961	AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys	2980
Db	9175	AATGGTTTCTTCTTAATCATGGGGCCATATACAGTATCAGTGTCTCTGTTTATAG	9234
Qy	2981	LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerPro	3000
Db	9235	CTCCATGGAAATTCATCAAGAGTGCTCTCCAAATGGCTCTCTGGAGTGGCAGCTCACCT	9294
Qy	3001	SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr	3020
Db	9295	TCTGCTCTGCTTGCAGATGTTCCACACCAAGTAATTAATGAAATATGAACTGTCAATGGGACA	9354
Qy	3021	AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly	3040
Db	9355	GAATTTGACTGTGAAAGGAGCCCGGATTCAGTGTCTTAAGGCTTCAAGCTTCTTAGGA	9414
Qy	3041	LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu	3060
Db	9415	CTTTCTGAAATCACCTGTGAAGCCGATGGCCAGTGGAGCTCTGGGTTCGCCCATCTGAA	9474
Qy	3061	HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer	3080
Db	9475	CACACTTCTGTGTGTTCTCTTCCAAATGATACCAAAATGCGTTTCATCATGAGAGACAGCTCT	9534
Qy	3081	TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer	3100
Db	9535	TGGAAAGGAAATGTGATTAATCTTACACTGCACTGCAAGTCTGGATATGTTCATACAGGCGAGTTCA	9594
Qy	3101	AspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpProValCysGluProLeu	3120
Db	9595	GATCTGATTTGTACAGAAAGGGGTATGGAGCCAGCTTATCCAGTCTGTGAGCCCTTG	9654
Qy	3121	SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr	3140
Db	9655	TCTGTGGGTCCCCACCGTCTGTGCGCAATGCGAGTGGCAACTGGAGAGGCCACCCACTAT	9714
Qy	3141	GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThr	3160
Db			
Db	9715	GAAGTGAAGTGAACCTCAGATGCTCGGAAGGTTTATACGATGATACAGATACAGATACA	9774
Qy	3161	PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys	3180
Db	9775	TTCACTGTTCAGAAAGATGCTCGTGGTCCCTCGTACAGAAATCTCTCGCAGTCTCTAAAAA	9834
Qy	3181	CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn	3200
Db	9835	TGTCTCTCTCCCGAAACATACACATATACCTTGTTCATGGGACGATTTTCAGTGTGAT	9894
Qy	3201	ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal	3220
Db	9895	AGGCAAGTTTCTGTTCATGTGCAGAAAGGTATACCTTTGAGGGAGTTAATACATATCAGTA	9954
Qy	3221	CysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSer	3240
Db	9955	TGTCACTTGTATGGAACCTCGGAGCCACCAATCTCCGATGAATCTTTCAGTCCAGTTTCT	10014
Qy	3241	CysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGlu	3260
Db	10015	TGTGGAAACCTGGAAGTCCAGAACATGGATTTGTGTGGCAGTAAATACACTTTTGAA	10074
Qy	3261	SerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgVal	3280
Db	10075	AGCACAAATTTATTCAGTGTGAGCTTGGCTATGAACCTAGAGGGGAACACAGGACGTGTC	10134
Qy	3281	CysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGlu	3300
Db	10135	TGCCAGGAGACAGACAGTGGAGTGGGGGTGGCAATATGCAAGAGACACAGGTGTGAA	10194
Qy	3301	ThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsn	3320
Db	10195	ACTCCACTTGAATTTCTCAATGGGAAGCTGCAATTTGAAACAGGACACTTGGACCCAC	10254
Qy	3321	ValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThr	3340
Db	10255	GTGGTATATTTCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCACACTGCACA	10314
Qy	3341	GluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValPro	3360
Db	10315	GAATAAGGAACCTGGAGCCACCAGTCCCTCTCTGCAAAACCAATCCATCCCTCTTCTCT	10374
Qy	3361	PheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnVal	3380
Db	10375	TTTGTGATTTCCCGAATGCTCTGCTGTCTGAAGAAGAGTTTATGTTGATCAGATGTG	10434
Qy	3381	SerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnPro	3400
Db	10435	TCCATCAATGTAGGGAAGGTTTCTGCTGAGGGCCACGGCATCATTTACTCTGCAACCCC	10494
Qy	3401	AspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAla	3420
Db	10495	GACGACGCTGGACACAGACAGCCGCAATGTGNAAAAAATCTCATGTGTGTCACACAGCT	10554
Qy	3421	HisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr	3440
Db	10555	CAGTAGAAAAATGCATTTGCTCGAGGGCTACATTTCAATATGGAGACATGATCACCTAC	10614
Qy	3441	SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly	3460
Db	10615	TCATGTTACAGTGCATACATGTTGGAGGTTTCTGCTGAGGAGTGTGTTCTTTAGAAAAATGGA	10674
Qy	3461	ThrTrpThrSerProPheCysArgAlaValCysArgPheProCysGlnAsnGlyGly	3480
Db	10675	ACATGACATCACCTCTATTTGACAGAGCTGTCTGTGATTTCCATGTGAGATGGGGGC	10734
Qy	3481	IleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuGlu	3500
Db	10735	ATCTGCCAGCCCAATGCTTCTCTGTCAGAGGGCTGGATGGGGGCTCTCTGTGAA	10794
Qy	3501	GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys	3520
Db	10795	GAACCAATCTGCATTTCTTCTCTGTCGAACGGAGGTGCTGCTGTGTGGGCCCTTACCAGTGT	10854

QY 3521 AspCysProGlyThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
 DB 10855 GACTGCCCGCTGCTGGAGCGGGTCTCGCTGATACAGCTTTTCCAGTCTCCCTGC 10914
 QY 3541 LeuAsnGlyLysCysValArgProAsnArgCysHisCysLeuSerSerThrGly 3560
 DB 10915 TTAATGGTGGAATGTGTAGACCAACCGATGCTCACTGTCTTTCTTCTGGACGGGA 10974
 QY 3561 HisAsnCysSerArgLysArgArgThrGlyPhe 3571
 DB 10975 CATAACTGTTCCAGGAAAGAGGACTGGGTTT 11007

RESULT 3

ABT08491
 ID ABT08491 standard; cDNA; 11152 BP.

AC ABT08491;
 XX
 XX
 XX
 XX 28-NOV-2002 (first entry)
 XX Human novel protein NOV2b coding sequence SEQ ID NO: 7.
 XX Human; NOVX; single nucleotide polymorphism; SNP; anti-HIV; cytostatic;
 KW antiarteriosclerotic; antidiabetic; antiasthmatic; antiinflammatory;
 KW haemostatic; hypotensive; neuroprotective; anorectic; nootropic;
 KW antipressant; immunosuppressive; antibacterial; antiparasitic;
 KW virucide; tranquilizer; anticonvulsant; osteopathic; analgesic;
 KW antiparkinsonian; dermatological; antiinfertility; cerebroprotective;
 KW antiaddictive; gene; ss.
 XX Homo sapiens.

OS WO200259315-A2.

XX 01-AUG-2002.

XX 19-DEC-2001; 2001WO-US050076.

XX 19-DEC-2000; 2000US-0256619P.

PR 19-JAN-2001; 2001US-0262959P.

PR 28-FEB-2001; 2001US-0272408P.

PR 20-APR-2001; 2001US-0285189P.

PR 26-JUL-2001; 2001US-0308039P.

PR 09-AUG-2001; 2001US-0311266P.

XX (CURA-) CURAGEN CORP.

XX Shimketa RA, Patturajan M, Vernet CAM, Casman SJ, Malyankar U;
 PI Shenoy S, Spytek KA, Gangolli B, Miller C, Boldog P, Li L;
 PI Taupier RJ, Kehuda R, Smithson G, Zerhusen BD, Liu X, Colman SD;
 PI Tchernev V, Si J, Edinger S, Stone D, Sciore P, Millet I;
 PI Rothenberg M;

XX WPI; 2002-666903/71.

DR P-PSDB; ABJ10590.

XX New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders e.g.
 PT diabetes, Crohn's disease, atherosclerosis, cancer, Huntington's disease
 PT or Alzheimer's disease.

XX Claim 8; Page 43-45; 163pp; English.

XX The present invention provides the protein and coding sequences of
 CC several novel human proteins, designated NOVX. These can be used in the
 CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-
 CC Lindau syndrome, Alzheimer's disease, stroke, tuberosus sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,

CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,
 CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or
 CC graft-versus-host disease. The present sequence is a coding sequence of
 CC the invention

XX SQ Sequence 11152 BP; 2951 A; 2606 C; 2711 G; 2884 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 11152
 Score: 19616.00 Matches: 3510
 Percent Similarity: 98.82% Conservative: 16
 Best Local Similarity: 98.37% Mismatches: 40
 Query Match: 98.21% Indels: 2
 DB: 6 Gaps: 1

US-09-977-053-4 (1-3571) x ABT08491 (1-11152)

QY 4 ArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThrPheGlnGln 23
 DB 83 AGAATTTGCGCGGCTTGTGGGCTGCGGCTCGCTTCGCGGCTGGCGACCTTTTCAGCAG 142

QY 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
 DB 143 ATGTCCCGCTGGCGCAATTTTCAGCTTCGCTCTTCCCGAGACCGCGCGGGGCCCCC 202

QY 44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
 DB 203 GGGAGTATCCCGCGCGCGCTCTTGGCGACGAGCGCGGGGAGCAGAGTGAGCGG 262

QY 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
 DB 263 CTGGCCAGGCGTTTC-----CGCGTGGCGTGTCTGCGGAGCTCAGCGAGCGCTGGAG 316

QY 84 LeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArgSerGluLeu 103
 DB 317 CTGTCTTCTCTGGTGGAGATTCGTCCAGCGTGGCGGAGTCAATCTCCGAGCGAGCTC 376

QY 104 MetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
 DB 377 ATGTTCGTCGCAAGCTGCTGTCCGCTTCCCGTGGTCCCGCCAGCGCGCGTGGCC 436

QY 124 IleValThrPheSerSerLysAsnTrpValValProArgValAspTrpIleSerThrArg 143
 DB 437 ATCTGACCTTCTCGTCCAGACTAGCTGGTGGCGGCGTGCATTTACATCTCCACCGC 496

QY 144 ArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIleSerTrpArg 163
 DB 497 CGCGCGCGCCAGCAAGTGGCGCTGCTCTCCAGAGATCCCTGCGCATCTCTTACCGA 556

QY 164 GlyGlyGlyThrTrpThrLysGlyValAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183
 DB 557 GGTGGCGGCGACCTACACCAAGGGCGCTTCCAGCAGCGCGCGCAAAATTTCTTTCATGCT 616

QY 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTrpSerAsnGlyGlyAsp 203
 DB 617 AGAGAAACTCAACAAAGTTGTATTTCTCATCTGATGATATTTCCAAATGGGGGAGAG 676

QY 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
 DB 677 CCTAGACCAATTTGAGCGTCACTCGAGATTCAGAGTGGAGATCTTCACTTTTGGCATA 736

QY 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
 DB 737 TGGCAAGGAACATTCGAGAGCTGATGACATGGCTTCCACCCCAAGAGGAGCACTGT 796

QY 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeuHisGluAsp 263
 DB 797 TACCTGTACACAGTTTTGAAGATTTTGAAGCTTTTAGCTCGCGCGGCAATTCATGAAGT 856

QY 264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrpLeuCysAspGlu 283
 DB 857 CTACCTCTTGGAGTTTTATTTCAGATGATATGTGCTCACTCTCAATCTTTGTGATGAG 916

QY 284 GlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHisPhe 303
DB 917 GGCAGGAGCTGCTGTGACCGAAATGGGAAGCTGCAAAATGTGGGACACACACAGGCCATTTT 976
QY 304 GluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThrAlaCys 323
DB 977 GAGTGCATCTGTGAAAGGGGTATACGGGAAGGTCTGCAGTATGAAATGCACAGCTTGC 1036
QY 324 ProSerGlyThrTyrLysProGluGlySerProGlyGlyLysSerSerCysIleProCys 343
DB 1037 CCATCGGGGACATACAAACCTGAAGCCTCAACAGGAGGAATCAGCAGTTGCAATTCATGT 1096
QY 344 ProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysValCysArg 363
DB 1097 CCGGATGAAAAATCACACCTCTCCACCTCGGAAGCACATCCCTCGAAGACTGTGTGTGCGA 1156
QY 364 GluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
DB 1157 GAGGGATACAGGCATCTGGCCAGACTGTGAACTTGTGCATCTGCCCTGCCCTGAGCCCT 1216
QY 384 ProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaIaCysGly 403
DB 1217 CCGGAAATGGTTACTTTATTCACAAACACTTGCACCAACCACTTCAATGCAGCCTGTGG 1276
QY 404 ValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeuProAsn 423
DB 1277 GTCCGATGTCAACCTGGATTTGATCTGTGGGAAGCAGCATCATCTTATGTCTACCCAA 1336
QY 424 GlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArgGln 443
DB 1337 GGTGTGTGTCGGTTTCAGAGACTACTGCAGATGAAACATGTCTCATCTCCGCCAG 1396
QY 444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThrCysLeu 463
DB 1397 CCGAAACATGGCCACATCAGCTGTCTACAAGGGAATGTTATATAAGACAAACATGTTG 1456
QY 464 ValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnValAsn 483
DB 1457 GTTGCCTGTGTGAAGGGTACAGACTAGAGAGGAGTGTAAGCTTACTTGTCAAGGAAC 1516
QY 484 SerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
DB 1517 AGCCAGTGGATGGCCAGAACCCCGTGTGTGGAGGCCACTGTCTCCACCTTTCAGATG 1576
QY 504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
DB 1577 CCGAAAGATGTTCATCATATCCCCCACTGTGGCAAGCAGCCAGCAAAATTTGGGACG 1636
QY 524 IleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
DB 1637 ATCTGCTATGTAAAGTTGGCCCAAGGGTTCAATTTATCTGGAGTCAAGAAATGCTGAGA 1696
QY 544 CysThrThrSerGlyLysTyrAsnValGlyValGlnAlaValCysLysAspValGlu 563
DB 1697 TGTAACACTTCTGGAAATGGAATGTGGAGTTCCAGCAGCTGTGTGTAAGACGTGGAG 1756
QY 564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnAspSer 583
DB 1757 GCTCTCTCAATCAACTGTCTAAGGACATAGAGGCTAAGACTCTGGAAACAGCAAGATTCT 1816
QY 584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysValSerVal 603
DB 1817 GCCAAATGTGTACCTGGCAGATTCCACAGCTTAAGACAACCTCTGGTGAAGGGTGCAGTC 1876
QY 604 HisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAlaIleVal 623
DB 1877 CCGGTTTCATCCACTTTCACCCCACTTACCTTTTCCCAATTCAGAGATTGTGTATCGTA 1936
QY 624 TyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLysValIle 643
DB 1937 TACAGGCCAACTGACCTATCCGGCAACAGGCCAGCTGCAATTTTCCATATCAAGGTTATT 1996
QY 644 AspAlaGluProProValIleAspTyrCysArgSerProProValGlnValSerGlu 663

DB 1997 GATGCAGAACCCCTCTCATAGACTGTGTGAGATCTCCACTCCCGTCCAGGTTCTCGAG 2056
QY 664 LysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu 683
DB 2057 AAGTACATGTCGCAAGCTGGGATGAGCTCAGTTCTCAGACAACCTCAGGGGCTGAATTG 2116
QY 684 ValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGln 703
DB 2117 GTCAATTACCAGAGTATACACAAGGAGACCTTTTCCCTCAAGGGGAGACTATAGTACAG 2176
QY 704 TyrThrAlaThrAspProSerGlyAsnAspArgThrCysAspIleHisIleValIleLys 723
DB 2177 TATACAGCCACTGACCCCTCAGGTAAACAGGATATGTGATATCCATATTGTCTAGAA 2236
QY 724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
DB 2237 GGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATATGCACCTCCAGAT 2296
QY 744 AsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySer 763
DB 2297 AATACTGGAGTCAACTGTACATTAACCTTGTCTGGAGGGCTACGATTTTCACAGAAGGGTCT 2356
QY 764 ThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrThrGlu 783
DB 2357 ACTGACAAGTATTATTGTGCTTATGAAGATGGGCTCTGGAAACCAACATATACCACTGAA 2416
QY 784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
DB 2417 TGSCCAGACTGTGCCAAAAAAGCTTTTGCAAAACCAAGGTTCAAGTCTCTTGAGATGTTT 2476
QY 804 TyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGlu 823
DB 2477 TACAAAGCAGCTGCTGTGTGATGACTCAGATCTCATGAAGAAGTTTCTGAAGCAATTTGAG 2536
QY 824 ThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
DB 2537 ACGACCTCGGAAATATGTTGCCATCATTTTGTGTGATGATGCAGAGGACATTTGACTGCGA 2596
QY 844 LeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGly 863
DB 2597 CTGGAGGAGAACCTGCACCAAAAAATATGCTAGAAATATATATATGATATGAAAAATGGC 2656
QY 864 PheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAsp 883
DB 2657 TTTGCAATTTGGTCCAGTGGCTGGGGTGCAGCTAAATAGGCTGGATTAATCTTACAGATGAC 2716
QY 884 PheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIle 903
DB 2717 TTCTTGACACTGTGCAAGAAACAGCCCAAGCATCGGCATGCGCAATGCCAAGTCTCAGCGATT 2776
QY 904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
DB 2777 AAAAGAAGTGGCCCATTTATCTGACTATATAAATTAAGTTAAATTTTAAATCATCACAGTAGT 2836
QY 924 ValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnAlaGluLeu 943
DB 2837 GTCCCATTTACCCCATGAAAGAAATGATPACCTTGAATGGGAAATCAGACACAGACTCTCT 2896
QY 944 GlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyr 963
DB 2897 CAGACATTTGGAAACTATCAAAATAAATCGAAAGGACTCTCAACAAAGACCCCATGTAT 2956
QY 964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLys 983
DB 2957 TCCCTTTTCAGCTTGCATCAGAAATACTTATATAGCCACAGCAATTCATTTAGAAACAAAAAG 3016
QY 984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003
DB 3017 GCTTCCCTCTTCGACAGACAGGCTCAGTGTAGAGGGGCTATGTGTGTCAATGGCCT 3076
QY 1004 LeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyr 1023

Db 3077 TTGGGACCTATTATTAATCTGGNACATTTTCCACTGTGAAAGCTGCCGGATCGGATCCTAT 3136
Qy 1024 GlnAspGluGluGlnLeuGluCysLysLeuCysProSerSerGlyMetTyrThrGluTyr 1043
Db 3137 CAAGATGAAGAAGGGCAACTTGAGTCAAGCTTGGCCCTCTGGGATGTACACGGAATAT 3196
Qy 1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
Db 3197 ATCCATTTCAAGAACATCTCTGATTTAAAGCTCAGTGTAAACCAAGCCACTTACTCATAC 3256
Qy 1064 SerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPheGlySer 1083
Db 3257 AGTGGACTTGAGACTTGTGAATCGTCTCCACTGGGCACCTATCAGCCAAAATTTGGTTC 3316
Qy 1084 ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle 1103
Db 3317 CGGAGCTGCTCTCGTGTCCAGAAAACACCTCACTGTGAAAAGAGAGCCGTGAACATT 3376
Qy 1104 SerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCys 1123
Db 3377 TCTGCATGTGGAGTTCCTTGTCCAGAAAGGAAAATTTCTGCGTTCCTGGGTTAATGCCCTGT 3436
Qy 1124 HisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys 1143
Db 3437 CACCCATGCTCTCGTGACTATTACCAACCTAATGCGAGGAAGCCCTTCTGCCCTGGCTGT 3496
Qy 1144 ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPhe 1163
Db 3497 CCTTTTATGGAACTACCCATTCCTCGTGTTCAGATCCATCAGAAATGTTCAAGTTTT 3556
Qy 1164 SerSerThrPheSerAlaGluGluSerValValProProAlaSerLeuGlyHisIle 1183
Db 3557 AGTTCAACTTTCTCAGCGGCAGAGGAAGTGTGTGCCCCCTGCTCTCTTGACACATATT 3616
Qy 1184 LysLysArgHisGluLeuSerSerGlnValPheHisGluCysPhePheAsnProCysHis 1203
Db 3617 AAAAAGGCATGAAATCAGCATCAGGCAGTCAATGATGCTTCTTTAACCTTGCCAC 3676
Qy 1204 AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr 1223
Db 3677 AATAGTGAACCTGCCAGCACTTGGCGTGTGTTATGTTCTCTCTCCACTTGGATAT 3736
Qy 1224 ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsn 1243
Db 3737 ACAGGTTTAAAGTGTGAAAACAGACATCGATGAGTGCAGCCCTGCTTGCCTCAACAAT 3796
Qy 1244 GlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGly 1263
Db 3797 GGAGTTTGTAAAGCTAGTTGGGGAAATTCATTTGTGAGTGCCCATCAGGTTACACAGGT 3856
Qy 1264 GlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIle 1283
Db 3857 AAGCACTGTGAATTGAACATCAATGAATGTCACTTAATCCATGTAGAAATCAGGCCACC 3916
Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
Db 3917 TGTGTGGATGAATTAATTCATACAGTTGTAAATGTTCAGCCAGGATTTTCAGGCAAAAGG 3976
Qy 1304 CysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCysGlu 1323
Db 3977 TGTGAACAGGATGATATCACTCAGTGTATTATTAATCACTTATATGATGATCTGTGAA 4036
Qy 1324 AspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCysGly 1343
Db 4037 GACCAGGTTGGGGATTCCTTGTGCAAAATGCCACCTCGATTTTTTGGGTACCCGATGTGA 4096
Qy 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363
Db 4097 AAGAAGCTTCGATGATGTCTCAGTCAGCCATGCAAAAATGGAGCTACTCTGAAGACGGT 4156
Qy 1364 AlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeuAsn 1383
Db 4157 GCCAATAGCTTCAGGTGCTTGTGTGAGCTGGCTTCACAGATCACACTGTGAATTGAAAC 4216

Qy 1384 IleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeuAsn 1403
Db 4217 ATCAATGAATGTTCAGTCTAATTCATGTAGAAATCAGGCCACCTGTGTGGATGAATTAAT 4276
Qy 1404 SerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGluGlnSer 1423
Db 4277 TCATACAGTTGTAAATGTTCAGCCAGGATTTTCAGGCAAAAGGTGTGAAACAGACAGTCT 4336
Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
Db 4337 ACAGGCTTTAACTCGATTTTGAAGTTTTCAGGATCTATGCGATATGTCATCTAGATGGC 4396
Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSerSerAspMet 1463
Db 4397 ATGCTCCCATCTCTCCATGCTCTAAACCTGTACCTTCTGGATGAATCTCTGACGACATG 4456
Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
Db 4457 AACTATGGAAACCAATCTCTATGCGATTTGATNACGGCAGCGCAATACCTTGTCTCTG 4516
Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsnCysPro 1503
Db 4517 ACTGATTATAACGGTGGTTCCTTATGTGATGCGAGGAAAGATAACAACCTGCTCC 4576
Qy 1504 SerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSerAlaAsnGlyIle 1523
Db 4577 TCGGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGACAAGTACTTGGTGGAGCC 4636
Qy 1524 TrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerValGlyLeu 1543
Db 4637 TGGAGGTCTATATAAATGGGAATTAATGACGGTGTACTGGCTCTCCATTGGCAAA 4696
Qy 1544 ProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluAspLysLysGlyGlyGly 1563
Db 4697 GCCATACCTGTCGGTGTGATTAGTTCTTGGGCAAGAGCAACAAAAGGAGAGGG 4756
Qy 1564 PheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAspTyrVal 1583
Db 4757 TTCAACCCGGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGACTATGTC 4816
Qy 1584 LeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGly 1603
Db 4817 CTGTCTCCACAGCGGTGAAGTCACTGGCTACTCTCTGCCCGAGAGAACTCAGTAAAGGA 4876
Qy 1604 AsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLysIleAspSer 1623
Db 4877 AACGTGTAGCATGGCTGATTTCTTGTAGGAATTTGTGGGAAAGTGAAGATCGATTCT 4936
Qy 1624 LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr 1643
Db 4937 AAGAGCATATTTTGTCTGATTTGCCACGCTTGGAGGGTCAGTGCCTCATCTGAGAACT 4996
Qy 1644 AlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGln 1663
Db 4997 GCATCTGAAGATTTAAACCCAGGTTCCAAAGTCAATCTGTTCTGTGAACAGGCTTCAG 5056
Qy 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyIleThrGlnProLeuPro 1683
Db 5057 CTGCTGGGAAACCTGTGTGAGTCTGTCTGAATCAAGACAGTGGACACCAACCTCCCTCC 5116
Qy 1684 HisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAla 1703
Db 5117 CACTGTGAACGCAATTCGCTGTGGGGTGCCACCTCTCTTGGAGAAATGGCTTCCATTGACGC 5176
Qy 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu 1723
Db 5177 GATGACTTCTATGCTGGCAGCAGCAGTAACCTACCAAGTGCACAAATGGCTACTATCTATG 5236
Qy 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerProSerCys 1743
Db 5237 GGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGGCTTTCACCATCTGTC 5296

Qy	1744	LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn	1766
Db	5297	TTAGATGTGCGATGAGTGTGCGAGTTGGATCAGATTGTAGTGAGCATGCTTCTTGCTGAAC	5356
Qy	1764	ValAspGlySerThrIleCysSerCysValProProTyrThrGlyAspGlyIysAsnCys	1783
Db	5357	GTAGATGGATCTCTACATATGTTTCATGTGCCACCGTACACAGAGATGGGAAAACTGT	5416
Qy	1784	AlaGluProIleLysCysLysAlaProGluValAsnProGluAsnGlyHisSerSerGlyGlu	1803
Db	5417	GCAGAACCTATAAATGTAAAGCTCCAGGAAATCCGGAAATGGCCACTCTCTCAGGTGAG	5476
Qy	1804	IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly	1823
Db	5477	ATTATTATACAGTAGGTGCCGAAGTCACATTTTCGTGTTCAGGAAGATACACAGTTGATGGGA	5536
Qy	1824	ValThrIlysIleThrCysLysLeuGluSerGlyGluTrpAsnHisIleuIleProTyrCysLys	1843
Db	5537	GTAACCAAAATACATATGTTTGGAGTCTCGGAATGGGAATCATCTAAATACCATATTGTAA	5596
Qy	1844	AlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGluLeuAlaPhe	1863
Db	5597	GCTGTTTTCATGTGTGAACCGGCTATTTCAGAAATGGTTGCATTTAGGAGTTAGCAATTT	5656
Qy	1864	ThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLys	1883
Db	5657	ACTTTTGGCAGCAAGTGACATATAGGTGTAATAAGATATATCTCTGGCGGGTGATAA	5716
Qy	1884	GluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCysGluProVal	1903
Db	5717	GAATCATCTCTGTCTGTAAACAGTCTCTGGAGTCAATTCCTCTCTGTGTGAACCAAGT	5776
Qy	1904	LysCysSerSerProGluAsnIleAsnAsnGlyIysTrpIleLeuSerGlyLeuThrTyr	1923
Db	5777	AAGTGTCTCTAGTCCGAAAAATATAAATAATGAANAATATATTTTGATGGGCTTACTAC	5836
Qy	1924	LeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIle	1943
Db	5837	CTTTCTACTGCATCATATTCATGCGATACAGATACAGCTTACAGGCGCTTCCATTATT	5896
Qy	1944	GluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCys	1963
Db	5897	GAATGCACGGCTTCTCGGCATCTCGGACAGAGCGCACCTGCCTGTCACTCGCTCTTCTGT	5956
Qy	1964	GlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsn	1983
Db	5957	GGAGAACCTCTGCCATCAAAAGATGCTGTCTANTAGGGGAATPACTTCACTTTTCAGAAC	6016
Qy	1984	ThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCys	2003
Db	6017	ACCGTCACTTACACTTGCAGGAAGGCTATACTCTCTGCTGTGTCACACCATTTGAATGC	6076
Qy	2004	LeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCysAspGlu	2023
Db	6077	CTGGCCGACGGCAAGTGGAGTAGAAGTAGCACAGAGTGCCTGGCTGTCTCTCTGTGATGAG	6136
Qy	2024	ProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAspIleAla	2043
Db	6137	CCACCCATTGTGGACACGCTCTCCAGAGACTGCCCATGGCTCTTTTGAGAGCATTTGCA	6196
Qy	2044	PheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAla	2063
Db	6197	TTCTACTACTGCTCTGATGGTTACAGCTACGACACAATTTCCAGCTTCTCTGCAATGCC	6256
Qy	2064	GlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAlaHisPheCys	2083
Db	6257	CAGGCAAGTGGGTACCCCAAGAGGTCAAGACATGCCCCGTGGTGTGTATAGTCTCAATTTCTGT	6316
Qy	2084	GlnLysProProSerValSerTyrSerIleLeuGluSerValSerLysAlaLysPheAla	2103
Db	6317	GAATAAACCTCCATCGGTTTCTCTATAGCATCTTTGGAATCTGTGAGCAAGCAAAATTTGCA	6376
Qy	2104	AlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLys	2123

Db	6377	GCTGGCTCAGTGTGGAGCTTTAAATGCATGGAAGGCTTTGTACTGAAACCACTCAGCAAG	6436
Qy	2124	IleGluCysMetArgGlyGlyGlnTnpAsnProSerProMetSerIleGlnCysIlePro	2143
Db	6437	ATTGAAATGATAGAGAGTGGGAGTGGAGACCTTTCCCCANATGTCATCCAGTGCATCCCT	6496
Qy	2144	ValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSer	2163
Db	6497	GTGCGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCAAACTACAGT	6556
Qy	2164	PheGlyAlaMetValIalatySerCysAsnLysGlyPheTyrIleIysGlyGluLysLys	2183
Db	6557	TTTGGAGCCATGGTGGCTTACAGCTGCACCAAGGGTTCACATCAAAAGGGGAAAGAAG	6616
Qy	2184	SerThrCysGluAlaThrGlyGlnTnpSerSerProIleProThrCysHisProValSer	2203
Db	6617	AGCACCTGCCAGGCCACAGGGCAGTGGAGTAGTCTATACCGACGTGCCACCCGGTATCT	6676
Qy	2204	CysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePhe	2223
Db	6677	TGTGGTGNACCCTAAGGTTGAGATGGCTTTCTGGAGCATACAACATGGCAGGATCTTT	6736
Qy	2224	GluSerGluValArgTyrGlnCysAsnProGlyTyrIlySerValGlySerProValPhe	2243
Db	6737	GAGAGTGAAGTGGAGTATCAGTGTAAACCCGGCTATAAGTCAGTCGGAATGCTCTGATTT	6796
Qy	2244	ValCysGlnAlaAsnArgHisTnpHisSerGluSerProLeuMetCysValProLeuAsp	2263
Db	6797	GTCTGCCAGCCAAATGCCACTGGCAGTGAATCCCTCTGATGTGTCTCTCTCGAC	6856
Qy	2264	CysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnGlyValGly	2283
Db	6857	TGTGGAAAACCTCCCCCGATCCAGATGGCTTCAATGAAGGAGAAAACTTTGAAGTAGGG	6916
Qy	2284	SerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerThrThr	2303
Db	6917	TCCAAGGTTTCAGTTTTCGTAAATGAGGGTTATGAGCTTTGTGGTGACAGTCTCTGGACA	6976
Qy	2304	CysGlnLysSerGlyLysTnpAsnLysLysSerAsnProLysCysMetProAlaLysCys	2323
Db	6977	TGTGAGAAATCTGGCAAAATGGAAATAGAGTCAATCCAAAGTGCAATGCTGCGCAAGTGC	7036
Qy	2324	ProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGly	2343
Db	7037	CCAGAGCGGCCCTCTTGGAAAACCAAGCTAGTATTAAAGGAGTTGACCCAGGAGTAGGA	7096
Qy	2344	ValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCys	2363
Db	7097	GTGTGACATTTTCTGTATAAGAGAGGCATGCTCTGCAAGGGCCCTCTGTCTCGAATGC	7156
Qy	2364	LeuProSerGlnTnpAsnAspSerPheProValCysLysIleValLeuCysThrPro	2383
Db	7157	TTGCCATCCAGCAATGGAAATGACTCTTTCCTGTTGTAGATTTGTTCTTGATCCCCA	7216
Qy	2384	ProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGlySerThr	2403
Db	7217	CCTCCCCCTAATTCCTTTGGTGTGCCCATTCCTCTCTCTCTCATTTTGGAGAGTACT	7276
Qy	2404	ValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGln	2423
Db	7277	GTCAAGTATTTCTGTGTAGGTGGTTTTTCTTAAGAGGAAATTTCTACCAACCTCTGCCAA	7336
Qy	2424	ProAspGlyThrTnpSerProLeuProGluCysValProValGluCysProGlnPro	2443
Db	7337	CCTGATGGCACCTGGAGCTCTCCACTGCCAGAAATGTTCACAGTAGAATGTCCCCAACCT	7396
Qy	2444	GluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeu	2463
Db	7397	GAGGAAATCCCCAATGGAATCAATTGATGTGCAAGGCTTTGCCCTATCTCAGCAGACTCTC	7456
Qy	2464	TyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsn	2483

Db 7457 TATACCTCGAACCGCTTTGAAATTTGGTGGGAAATATACCAACCCCTTTGTGGAGAAAT 7516
Qy 2484 GlyHisTyrLeuGlyGlyLysProThrCysLysAlaIleLeuCysLeuLysProLysGlu 2503
Db 7517 GGTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTGAGTGCCTTGAACCAAGGAG 7576
Qy 2504 IleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSer 2523
Db 7577 ATTTTGAATGGCAATCTCTTACACGGACCTACACTATGACAGACCGTTACTTACTCT 7636
Qy 2524 CysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAsp 2543
Db 7637 TGCAACCGAGGCTTTCGGCTTCGAAGTCCAGTGCCTTGACCTGTGTTAGAGACAGGTGAT 7696
Qy 2544 TrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGlu 2563
Db 7697 TGGGATGTAGATGCCCATCTTGTCAATGCCAATCACTGTGATTTCCCAACCAATGAA 7756
Qy 2564 AsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleTyrSerCysPhe 2583
Db 7757 AATGGTTTTGTAGAGTGCAGATTACAGCTATGGTGCCTATCATCTACAGTTGCTTC 7816
Qy 2584 ProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysLeuGluSerGlyTrpSerSer 2603
Db 7817 CTTGGGTTTTCAGGTGGCTGTCAATGCATGCACACCTGTGAAGATCAGGATGGTCAAGT 7876
Qy 2604 SerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPheGlyAsp 2623
Db 7877 TCCATCCCAACATGATGCCATAGACTGTGGCTCTCCCTCATATAGATTTTGGAGAC 7936
Qy 2624 CysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspMetMetGluVal 2643
Db 7937 TGTACTAACTCAAGATGACCAAGATATTTTGGACAGAGACCAATGATGGAGTT 7996
Qy 2644 ProTyrValThrProHisProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsn 2663
Db 7997 CCATATGTGACTCTCACCTCTCTTATCATTTTGGGACAGTGGCTAAACCTGGGAAAT 8056
Qy 2664 ThrLysGluSerProIleThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyr 2683
Db 8057 ACAAGGAGTCTCTCTCTACACATTCATCAAACTTTCTGTATGGTACCATTGGTTTCTATC 8116
Qy 2684 ThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValIleLeuLysGlnGluAspGly 2703
Db 8117 ACTGTATCCAGATATGAACTCTGGGGAACCTGTGCTGATCTGCCAGGAGATGGA 8176
Qy 2704 ThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaPro 2723
Db 8177 ACTTGGATGGCAGTGCACCATCTCTCAATTTCAATTTGAATGTGACTTGCTACTGCTCT 8236
Qy 2724 GluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCys 2743
Db 8237 GAAATGGCTTTTGGCTTTTACAGAGCTAGCATGGGAAGTGTGTGACAGTATAGCTGT 8296
Qy 2744 LysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrp 2763
Db 8297 AAACCTGGACACATCTACAGGCTCTGACTTAAGGCTTTGTCTAGAGATAGAAAGTGG 8356
Qy 2764 SerGlyAlaSerProArgCysGluAlaIleSerCysLysLysPheProAsnProValMetAsn 2783
Db 8357 AGTGGTGCCTCCCGCTGTGAAGCAATTCATGCAAAAGCCAAATCCAGTCATGAAT 8416
Qy 2784 GlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspPro 2803
Db 8417 GGATCCATCAAGGAAGCAACTACATACCTGAGCACCTTGTACTATGAGTGTGACCCC 8476
Qy 2804 GlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspLysAsnTrpAspGlu 2823
Db 8477 GGATATGTCTGATGGCACTGAGAGAGAACATGCCAGGATGACAAACTGGGATGAG 8536
Qy 2824 AspGluProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyGln 2843
Db 8537 GATGAGCCCATTTGCAATCTCTGTGAGTGCAGTTTACCCCCCAGTCTCAGCCCAATGGCCAG 8596

Qy 2844 ValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPhe 2863
Db 8597 GTGAGGAGAGACGATACACATTTCCAAAAGAGATTGATATACATTTGCATATGAGGGTTC 8656
Qy 2864 LeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThr 2883
Db 8657 TTGCTTTGAGGAGCCAGGAGTCGGTTTGTCTTGGCAATGGAAAGTTGGAGTGGAGCCACT 8716
Qy 2884 ProAspCysValProValArgCysAlaThrProGlnLeuAlaAsnGlyValThrGlu 2903
Db 8717 CCCGACTGTGTGCTGTGCATGTGCCACCCGCCACACTGGCCCAATGGGGTGAAGGAA 8776
Qy 2904 GlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeu 2923
Db 8777 GGCTGTGACTATGGCTTCATGAAGAGATGAACATTCCTCTGTACAGGGGCTACATCTTG 8836
Qy 2924 HisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeu 2943
Db 8837 CACGGTGTCCAAAACCTCACCTGTCTAGTCAGATGGCAACTGGGATGCAGAGATTCCTCTC 8896
Qy 2944 CysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPhe 2963
Db 8897 TGTAAACCACTGCTGAGACCTCTCTGAAGATCTTGCCTCATGTGTTCCCTAAATGGTTTT 8956
Qy 2964 SerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGly 2983
Db 8957 TCTTTTATTCATGGGGCCATATACAGTATCAGTGTCTTTCTGTGTTATAGCTCATGGA 9016
Qy 2984 AsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeu 3003
Db 9017 AATTATCAAGAGGTGCCTCTCAATGGCTCTCGAGTGGCAGCTCACCTTCCTGCTG 9076
Qy 3004 ProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAsp 3023
Db 9077 CCTTGCAGATGTTTCCACACCAAGTAATTGAATATGGAATCTGCAATGGGACAGATTTTGAC 9136
Qy 3024 CysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLysSerGlu 3043
Db 9137 TGTGGAAAGGACGCCGATTCAGTGTCTTCAAGGCTTCAAGCTCTTAGGACTTTCTGAA 9196
Qy 3044 IleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSer 3063
Db 9197 ATCACCTGTGAAGCCGATGGCAGTGGAGCTCTGGGTTCCCCACTGTGAACACACTTCT 9256
Qy 3064 CysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGlu 3083
Db 9257 TGTGGTTCTCTTCCAAATGATACCAATGCGTTTCATCAGTGAAGCCAGCTCTTGAAGGAA 9316
Qy 3084 AsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIle 3103
Db 9317 AATGTGATTAATCTACAGTGCAGTCTGGATATGTCTACAGGACGATTCAGATCTGAT 9376
Qy 3104 CysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSerCysGly 3123
Db 9377 TGTACAGAGAAAGGGGTATGAGCCAGCTTATCCAGTCTGTGAGCCCTTGTCTCTGTTGG 9436
Qy 3124 SerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGlu 3143
Db 9437 TCCCAACCGTCTGTGCCCAATGCAAGTGGCAACTGGAGAGGCACACACTATGAAAGGAA 9496
Qy 3144 ValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrPheThrCys 3163
Db 9497 GTGAAACTCGATGTCTGGAGGTTATACGATGGATACAGATACATACATCACCTGT 9556
Qy 3164 GlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeu 3183
Db 9557 CAGAAAGATGGTGGCTGTGTTCCCTGAGAGAAATCTCTGACAGTCTTAAAAAATGTCTCTC 9616
Qy 3184 ProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnVal 3203
Db 9617 CCGGAAACATACACATATACTTGTATCGGGGACGATTTTCAAGTGTGAAATAGGCAAGTT 9676

XX Claim 8; Page 38-41; 363pp; English.

PS The present invention provides the protein and coding sequences of

CC several novel human proteins, designated NOVX. These can be used in the

CC treatment of diseases such as cancers, Hodgkin's disease, Von Hippel-

CC Lindau syndrome, Alzheimer's disease, stroke, tuberosclerosis, cerebral

CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral

CC palsy, epilepsy, Leach-Nyhan syndrome, multiple sclerosis, ataxia

CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,

CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,

CC infertility, atherosclerosis, hypertension, scleroderma, haemophilia,

CC diabetes, pancreatitis, autoimmune disease, asthma, arthritis,

CC immunodeficiencies, HIV, viral, bacterial or parasitic infections, or

CC graft-versus-host disease. The present sequence is a coding sequence of

CC the invention

XX

SQ Sequence 11158 BP; 2952 A; 2603 C; 2710 G; 2893 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	11158
Score:	19566.00	Matches:	3503
Percent Similarity:	98.57%	Conservative:	16
Best Local Similarity:	98.12%	Mismatches:	47
Query Match:	97.96%	Indels:	4
DB:	6	Gaps:	2

US-09-977-053-4 (1-3571) x ABT08490 (1-11158)

QY 4 ArgLeuAlaPheCysCysTyrGlyLeuAlaLeuValSerGlyTyrAlaThrPheGlnGln 23

DB 83 AGAATTTCGGCGGCTTCTGGGGTCTGGGCTCGTTTCGGGCTGGGCGACCTTTTCAGCAG 142

QY 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43

DB 143 ATGTCCCGCTCGCGCAATTTTCAGCTTCGGCTTCCTCCCGAGACCGCGCGCGCGCC 202

QY 44 GlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63

DB 203 GGGAGTATCCCGCGCGCGCTCTGGCGACGAGCGCGGGAGCAGAGTGGAGCGG 262

QY 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83

DB 263 CTGGCGCAGGGGTTT-----CGCGTGGCGCTGCTGGCGGAGCTCAGCGAGCGCTGGAG 316

QY 84 LeuValPheLeuValAspAspSerSerValGlyGluValAlaAsnPheArgSerGluLeu 103

DB 317 CTGTGCTTCTGTGTGATGATTCGTCCAGCGTGGCGCGAAGTCAACTTCCGCGAGCGCTC 376

QY 104 MetPheValArgLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123

DB 377 ATGTTCCTCCGCAAGCTGTCTGCGACTTCCCGTGTGTCGCCGCGCAGCGCGGTGGCC 436

QY 124 IleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIleSerThrArg 143

DB 437 ATCGTGACCTTCTGTCCAAAGTACGTGTGTGGTGGCGCGCGTGGATTACATCTCCACCGC 496

QY 144 ArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGlnIleProAlaIleSerTyrArg 163

DB 497 CGCGCGCGCGCAGCAAGTGGCGCTGCTCTCCCAAGAGATCCCTGCGCATCTCTCCCGA 556

QY 164 GlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeuLeuHisAla 183

DB 557 GGTGGCGGCGCCTTACACCAAGGGCGCTTCAGCAGACGGCGCGCAATTTCTTCTCATGCT 616

QY 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAsp 203

DB 617 AGAGAAACTCAACAAAGTTGTATTTCTCATCTCATCTCATCTCATCTCATCTCATCT 676

QY 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223

DB 677 CCTAGACCAATTGCGCGCTCAGTGGAGATTTTCAGGAGTGGAGATCTTCACTTTTGGCAT 736

QY 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243

DB 737 TGGCAGGAGAACATTCGAGAGCTGATGACATGGCTTCCACCCCAAGGAGGAGGACTGT 796

QY 244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArg-----ArgAlaLeuHis 261

DB 797 TACTGCTACACAGATTTTGAAGAAATTCGAGGCTTAGTCCGCTCTGTCTATATGTTATTT 856

QY 262 GluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeuCys 281

DB 857 GTAGATCTACCTTCTGGGAGTTTATTCAGAGATATGTTCCACCTGCTCATATCTTTGT 916

QY 282 AspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHisThrGly 301

DB 917 GATGAGCGCAGGACTGCTGTGACCGAATGGAGCTGCAATGTGGGAAACACACAGGC 976

QY 302 HisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThr 321

DB 977 CATTTGAGTGCATCTGTCAAAAGGGTATATACGGGAAAGGCTGCAAGTATGACTGCACA 1036

QY 322 AlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCysIle 341

DB 1037 GTTTGCCCATCGGGACATACAACTGAAGCTCACCAGGAGGAATCAGCAGTTCAT 1096

QY 342 ProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysVal 361

DB 1097 CAGTCTCTGATGAATATCACACTCTCCACTGGAGCAGCATCCCTGGAAGCTGTGTC 1156

QY 362 CysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeu 381

DB 1157 TGCAGAGAGGAGATACAGGCGACTCTGGCAGACCTGTGAAGTTGTTCACCTGCCCTG 1216

QY 382 LysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAla 401

DB 1217 AAGCTCCCGAAATGGTTACTTTATCCAAACACTTGCACACACCTTCAATGAGCC 1276

QY 402 CysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCysLeu 421

DB 1277 TGTGGGTCGATGTCACCTGGATTTCATCTTGTGGGAGCAGCATCATCTTATGCTA 1336

QY 422 ProGlnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysProHisLeu 441

DB 1337 CCCAATGTTGTGTGTCGGTTTCAGAGAGCTACTGCGAGAGTAAGAACATGCTCTCATCTC 1396

QY 442 ArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThr 461

DB 1397 CGCCAGCGCAACATGGCCCATCATCAGCTGTCTACAAGGAAATGTTATATAGACAA 1456

QY 462 CysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGln 481

DB 1457 TGTGTTGTTGCTGTGATGAAGGTTACAGGCTAGAGGCGAGTATGATGCTTACTTGTCAA 1516

QY 482 GlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThrPhe 501

DB 1517 GGAACACCGCAGTGGAGTGGCGCAACCCCGTGTGTGGAGCGGCTGTTTCCACTTT 1576

QY 502 GlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPhe 521

DB 1577 CAGATGCCCAAGATGTATCATATATCCCCCAAACTGTGGCAGCAGCAGCAGCAATTT 1636

QY 522 GlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMet 541

DB 1637 GGGAGCATCTGCTATGTAAGTTGGCGCAAGGGTTTCAATTTATCTGGAGTCAAGAAATG 1696

QY 542 LeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAsp 561

DB 1697 CTGAGATGATACCTTCTGGAAAAATGGAATGTGCGAGTTTCAAGGAGCTGTGTGTAAAG 1756

QY 562 ValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlnGln 581

DB 1757 GTGAGGCTCTCAATCACTGTCTAAGGACATAGAGGCTTAGAGCTCTGGAACAGCAA 1816

QY 582 AspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLysVal 601

1817 GATTCTGCCAAATGTTACCTGGCAGATTCACACAGCTAAAGACCACTCTGGTGAAGAGGTG 1876
602 SerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAspValAla 621
1877 TCAGTCCAGCTTCATCCAGCTTCACCCACCTTACTCTTTCCAAATGGAGATGTGCT 1936
622 IleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLys 641
1937 ATCGTATACAGGCAACTGACCTATCCGGCAACAGCCAGCTGCATTTTCCATATCAAG 1996
642 ValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGlnVal 661
1997 GTTATTGATGCAGACCACTCTCATAGACTGGTGAGATCTCCACCTCCCGTCCAGGTC 2056
662 SerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGlyVal 681
2057 TCGAGAGAGTACATCCCGGAGCTGGATGAGACCTCAGTTCTCAGACCACTCAGGTGCT 2116
682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThrIle 701
2117 GAATTCGTCAATACCAAGAGTCATACACAGGAGACCTTTTCCCTCAAGGGGAGACTATA 2176
702 ValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
2177 GTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATGTC 2236
722 IleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThr 741
2237 ATAAAGGTTCTCCCTGTGAATTCCTTCCACCTGTAAATGGGATTTTATATGCACT 2296
742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
2297 CCAGATAAATACTGGAGTCAACTGTACATTAATCTGCTGGAGGCTATGATTTTACAGAA 2356
762 GlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrPheThrProThrTyrThr 781
2357 GGGTCTACTGACAAAGTATTTATGCTGTATGAGATGGCGCTCTGGAAACCAACATATACC 2416
782 ThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
2417 ACTGAATGGCCAGACTGTGCAGTAAGCTTTTGCACACCGGTTCAAGTCTCTTGAG 2476
802 MetPheTyrLysAlaLysCysAspThrAspLeuMetLysLysPheSerGluAla 821
2477 ATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTCATGAAGAAGTTTCTGAAGCA 2536
822 PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
2537 TTTGAGACGACCTGGGAAAAATGGTCCCATCATTTTGTAGTATGACAGGACATTTGAC 2596
842 CysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
2597 TGCAGACTGAGGAGAACCTGACCAAAAATATTCCTAGATATATATATGACTATGAA 2656
862 AsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyr 881
2657 AATGGCTTTGCAATTTGCTCCAGTGGCTGGGTGTCAGCTAATAGGCTGGATTTACTCTTAC 2716
882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
2717 GATGACTTCTTGACACTGTGCAAGAAACAGCCCAAGCATCGGCAATGCCAGTCTCTCA 2776
902 ArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThr 921
2777 CGGATTAAGAGAGTGGCCCCATTATCTGACTATATAAATAAGTTAAATTTTAAACATCACA 2836
922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnArg 941
2837 GCTAGTGTGCATTTACCCGATGAAGAAATGATACCCCTGAAATGGGAAAAATCAGAACGA 2896
942 LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961

2897 CTCCTTCAGACATTTGGAAACTATCACAAAATAAACTGAAAAAGGACTCTCAACAAGAGCCCC 2956
962 MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
2957 ATGTATTCTTTTCAGCTTGCATCAGAAATATCTATAGCCGACAGCAATTCATTAGAAACA 3016
982 LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
3017 AAAAAGGCTTTCCCTCTTCAGACAGGCTCAGTGTGTGAGAGGGCGTATGTGTGTCAAT 3076
1002 CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
3077 TGCCTTTTGGAAACCTTATATATCTGAAACATTTTACCTGTGAAAGCTGCCGATCGGA 3136
1022 SerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyrThr 1041
3137 TCCTATCAAGATCAAGAGGGCAACTTGAAGTCAAGCTTTGCCCTCTGGGATGTACAG 3196
1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
3197 GAATATATCCATTTCAAGAAACATCTCTGATTTGAAAGCTCAGTGTAAACAGGACCTTAC 3256
1062 SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPhe 1081
3257 TCATCAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCACCTTATCAGCCAAAATTT 3316
1082 GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaVal 1101
3317 GGTTCGGGAGCTGCCTCTCGTGTCCAGAAAACACCTCAACTGTGAAAAGAGGAGCGCTG 3376
1102 AsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMet 1121
3377 AACTTTCTGCATGTGGAGTTCTTGTTCAGAGAGAAAATTTCTCGGTTTCTGGGTAAAG 3436
1122 ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeu 1141
3437 CCTGTGACCCATGTCCTGTGACTATTACCACTTAATCAGGAGAGGCTTCTGCGCTG 3496
1142 AlaCysProPheTyrGlyThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
3497 GCCTGTCCCTTTTATGGAACCTACCCCAATCGCTGTTCAGAGTGTGGTCCCTCTCTTGA 3556
1162 SerPheSerSerThrPheSerAlaAlaGluSerValValProProAlaSerLeuGly 1181
3557 AGTTTATGTTCAACTTTCTCAGCGGACAGAGAAAGTGTGGTCCCTCTCTCTTGA 3616
1182 HisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnPro 1201
3617 CATATTAAGAAGAGGCAATGAATCAGCAGTCAAGCAAGTCATGAATGCTTCTTTAACCT 3676
1202 CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
3677 TGCACAAATAGTGGAACTTCCAGCAACTTGGGCGTGGTTATGTTGTCTCTGTCCACTT 3736
1222 GlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
3737 GGATATACAGTTTAAAGTGTGAACAGACATCGATGAGTGCAGCCCACTGCTCTGCTC 3796
1242 AsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
3797 AACAACTGGAGTTTGTAAAGACCTAGTTTGGGAAATTCATTTGTGAGTGGCCATCAGTTAC 3856
1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLys 1281
3857 ACAGGTAAGCACTGTGAATTTGAACATCAATGAATGTTCAGTCTAATCTCCTAGTAAATCAG 3916
1282 GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301
3917 GCCACCTGTGTGGATGAATTAATTCATAGTTGTAAATGTAGCCAGGATTTTCAGGC 3976
1302 LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
3977 AAAAGGTGTGAACAGGATGTATCAACTCAGTGTGTTATTAATAACCTTAATAATGCAGTC 4036

1322 CysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArg 1341
Db TGTGAAGACCAAGGTGGGGGATCTTGTGCAAAATGCCACCTGGATTTTGGGTACCCGA 4096
1342 CysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValAlaThrCysLys 1361
Db TGTGGAAGAACCTGCATGAGTGTCTCAGTCAGCCATGCCAAATGGAGCTACCTGTAA 4156
1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
Db GACGCTGCCAATAGCTTCAGGTGCTGTGTGAGCTGGCTTTACAGGATCACACTGTGAA 4216
1382 LeuAsnIleAlaGlnCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
Db TTGAACATCAATGAATGTCAGTCAATCATGTAGAAATCAGGCCACCTGTGTGGATGAA 4276
1402 LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
Db TTAATTCATACAGTTGTAAATGTCCAGCAGATTTTCAGGCAAAAGGTGTGAACAGAA 4336
1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
Db CAGTCTACAGGCTTTAACTGGAATTTTGAAGTTCCTGAGCTATGGATATGTCAATGCTA 4396
1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAsp 1461
Db GATGGCATGTCTCCATCTCTCATCTTAACCTGTATCTTCTGGATGAATCTCTCTGAC 4456
1462 AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeu 1481
Db GACATGAACATATGNAACCAATCTCTATGCACTTGATTAAGCGCAGCGACACATACCTTG 4516
1482 LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIleThrAsn 1501
Db CTCCTGACTGATTAATACCGGTGGGTCTTTATGTGAATGGCAGGAAAGATTAACAAC 4576
1502 CysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrSerAlaAsn 1521
Db TGTCCCTCGGTGAATGATGGCAGATGGCATCATATTTGCATCCTTGGCAAGTACTGGT 4636
1522 GlyIleTrpLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerVal 1541
Db GGACCTGGAGGCTATATAAATGGGAATTAATCTGACGGTGGTACTGGCTCTCCAT 4696
1542 GlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLysGly 1561
Db GCCAAAGCCATACCTGGTGGCGGTGCATTAAGTCTTGGGCAAGAGCAAGACAAAAAGGA 4756
1562 GluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAsp 1581
Db GAGGGGTTCAACCCCGCTGAGTCTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGAC 4816
1582 TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer 1601
Db TATGTCCTGTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTGCCCCAGAGAACTCAGT 4876
1602 LysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLysIle 1621
Db AAAGAAACCGTGTAGCATGCCCTGATTTCTGTTCAGGAATGTGGGGAAAGTGAAGATC 4936
1622 AspSerLysSerIlePheCysSerAspCysProArgLeuGlyLysSerValProHisLeu 1641
Db GATTCTAAGAGCATATTTGTCTTCATATGCCACGCTTGGGGGGTTCAGTGCCTCATCTG 4996
1642 ArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly 1661
Db AGAACTGCATCTGAAGATTTAAACACAGGTTCAAAGTCAATCTGTCTGTGAACACGCG 5056
1662 PheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnPro 1681
Db TTCAGCTGGTGGGAACCCCTGTGTGACGTACTGTCTGAATCAAGGACAGTGGACAAACCA 5116

1682 LeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHis 1701
Db CTCCTCCCACTGTGAACGATTCGCTGTGGGGTGGCCACCTCTCTTGGAGAAATGGCTTCCAT 5176
1702 SerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyr 1721
Db TCAGCCGATGACTTCTATGCTGGCAGCAGTAACCTACCACTGCAACAAATGGCTACTAT 5236
1722 LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerPro 1741
Db CTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCTTCCACA 5296
1742 SerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCys 1761
Db TCCTGCTTAGATGTCAGTAGTGAGTGCAGTTCAGATTCAGATTGTAGTGAGCATGCTTCTTGC 5356
1762 LeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLys 1781
Db CTGAACGTAGATGGATCCTACATATGTTCCATGTTCCACCGCTACACAGGAGATGGGAA 5416
1782 AsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSerSer 1801
Db AACTGTGCAGAACCTTATAAATGTAAAGCTCCAGGAAATCCGGAAATGGCCACTCTCA 5476
1802 GlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeu 1821
Db GGTGAGATTATACAGTAGGTCCGAGTCACTTTTCTGTTCAGGAGGATACCAAGTTG 5536
1822 MetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIleProTyr 1841
Db ATGGAGTAACCAAAATCACATGTTTGGAGTCTGCAGAAATGGAATCATCTAATACCATAT 5596
1842 CysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGluLeu 1861
Db TGTAAAGCTGTTTCATGTGGTAAACCGGCTATTCCAGAAATGGTGTGCATTGAGGAGTTA 5556
1862 AlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGly 1881
Db GCATTTACTTTTGGCAGCAGTAAGTGCATATAGTGTATTAAGGATATATCTCTGGCCGCT 5716
1882 AspLysGluSerSerCysLeuAlaAsnSerTrpSerHisSerProProValCysGlu 1901
Db GATAAAGAAATCATCTGCTTGTAAACAGTCTTGGAGTCAATTCCTCTCTGTGTGAA 5776
1902 ProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeu 1921
Db CCAGTGAAGTGTCTAGTCCCGGAAATATAATAATAGAAATAATATTTTGTAGTGGCTTT 5836
1922 ThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer 1941
Db ACCTACCTTTCTACTGTCATCATATTCATGCCGATACAGGATACAGCTTACAGGGCCCTTCC 5896
1942 IleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeuVal 1961
Db ATTATTTGAATGACCGCTTCTGGCATCTGGGACAGAGCGCCACCTGCTGTCACTCTGTC 5956
1962 PheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPhe 1981
Db TTCTGTGGAGAACCAACCTGCCATCAAGATGCTGCATTAACGGGAAATTAACCTTCACTTTC 6016
1982 ArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIle 2001
Db AGGAACACCGTCACCTTACCTTGCAGAGAGGCTATACTTCTTGTGCTTCTTGACACCAT 6076
2002 GluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCys 2021
Db GAATGCTCTGGCGAGCGCAAGTGGAGTAGAGTACCAGCAGTGGCTGTGCTGTCTCTGT 6136
2022 AspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAsp 2041
Db GATGAGCCACCCATTTGTGGACCAACCCCTCTCAGAGACTGGCCATCGGCTCTTTGGAGAC 6196
2042 IleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCys 2061

Db	6197	ATGTGATTCTACTACTGCTCTGATGGTTACAGCTTAGCAGACAATTCACAGCTTCTCTGC	6256
Qy	2062	AsnAlaGlnGlyLysTrpValProGluGlyGlnAspMetProArgCysIleAlaHis	2081
Db	6257	AATGCCAGGGCAGTGGGTACCCCAAGAGGTACAGACATGCCCGTTGTATAGTCAAT	6316
Qy	2082	PheCysGluLysProProSerValSerTrpSerIleLeuGluSerValSerIysAlaLys	2101
Db	6317	TTCTGTGAAAAACCTCCATCGGTTTCTATAGCATCTTGGAACTCTGTGAGCAAGCAAAA	6376
Qy	2102	PheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSer	2121
Db	6377	TTTTCAGCTGGCTCAGTTGTGAGCTTTAAATGCAAGAGGCTTTGTACTGAACACCTCA	6436
Qy	2122	AlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGlnCys	2141
Db	6437	GCAAGATTGAAATGATGAGAGTGGGAGTGGAGCCCTTCCCATGCTCCATCCAGTGC	6496
Qy	2142	IleProValArgCysGlyGluProProSerIleMetAsnGlyTrpAlaSerGlySerAsn	2161
Db	6497	ATCCCTGTGGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCAAC	6556
Qy	2162	TyrSerPheGlyAlaMetValAlaLysTrpSerCysAsnLysGlyPheTrpIleLysGlyGlu	2181
Db	6557	TACAGTTTGGAGCCCAATGGTGTACAGCTGCACAAGGGGTTCACATCAAGGGGAA	6616
Qy	2182	LysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisPro	2201
Db	6617	AAGAAGAGCACCTGCGAGAGCCAGAGGAGTGGAGTAGTCTATATACCGAGCTGCCACCG	6676
Qy	2202	ValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrGlyArg	2221
Db	6677	GTATCTGTGGTGAACACCTAAGGTGTGAATGGCTTTCTGAGGACATACAACTGGCAGG	6736
Qy	2222	IlePheGluSerGluValArgTrpGlnCysAsnProGlyTrpLysSerValGlySerPro	2241
Db	6737	ATCTTTGAGAGTGAAGTGAAGTATCAGTGTAAACCCGGGCTATGAAGTCAAGTGGAGTCT	6796
Qy	2242	ValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValPro	2261
Db	6797	GTATTTGTCTGCCAAGCCATCGCCATCGCCACAGTGAATCCCTCTCTGATGTGTTCT	6856
Qy	2262	LeuAspCysGlyLysProProLysIleGlnAsnGlyPheMetLysGlyGluAsnPheGlu	2281
Db	6857	CTCGACTGTGAAAAACCTCCCGGATCCAGAATGGCTTCATGAAAGGAGAAAACTTTGAA	6916
Qy	2282	ValGlySerLysValGlnPhePheCysAsnGluGlyTrpGluLeuValGlyAspSerSer	2301
Db	6917	GTAGGGTCCAAGGTTCAAGTTTCTGTAAATGAGGGTTATGAGCTTGTGTGACAGTTCT	6976
Qy	2302	TrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAla	2321
Db	6977	TGGACATGTCAGAAATCTGCCAAATGGAATAAGAGTCAAAATCCAAAGTGCATGCCCTGCC	7036
Qy	2322	LysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrGlu	2341
Db	7037	AAGTGGCCAGAGCGCGCCCTCTTGGAAAAACAGCTAGTATTAAGAGGTTGACACCGAG	7096
Qy	2342	ValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerValLeu	2361
Db	7097	GTAGGAGTTGTGACATTTTCTGTAAATGAGGGTTATGAGCTTGTGTGACAGTTCT	7156
Qy	2362	LysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCys	2381
Db	7157	AAATGCTTCCCATCCAGCAATGGAATGACTCTTCCCTGTTTGTAAAGATTGTCTTTGT	7216
Qy	2382	ThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGly	2401
Db	7217	ACCCACCTCCCTTAATTTCTTTGGTGTGCCCAATTCCTTCTCTGCTCTCTCAATTTGA	7276
Qy	2402	SerThrValIleTrpSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeu	2421

Db	7277	AGTACTGTCAAGTATTTCTTGTAGTGGGTTTTTCTTAAGAGAAATTTCTACCAACCTC	7336
Qy	2422	CysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCysPro	2441
Db	7337	TGCCAACCTGATGGCACCTCTGAGCTCTCCACTGCCAGAAATGTGTCTCCAGTAGAATGTCCC	7396
Qy	2442	GlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTrpLeuSerThr	2461
Db	7397	CAACCTGAGGAATATCCCAATGGAATCATTTGATGTCAAGGCTTGCCTATCTCAGACACA	7456
Qy	2462	AlaLeuTrpThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGly	2481
Db	7457	GCTCTCTATACCTGACAGCCAGGCTTTGAATGTGTGGAAATTACTACCAACCTTTTGGGA	7516
Qy	2482	GluAsnGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeuLysPro	2501
Db	7517	GAATAATGGTCACTGGCTTGGAGGAAAAACAACATGTAAGCCATTTGAGTGCCTGAAACCC	7576
Qy	2502	LysGluIleLeuAsnGlyLysPheSerTrpThrAspLeuHisTrpGlyGlnThrValThr	2521
Db	7577	AAGGAGATTTTGAATGGCAAAATTTCTTTACACGAGCTTACATATGGACAGACCGTTACC	7636
Qy	2522	TyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThr	2541
Db	7637	TACTCTTGCACCGAGGCTTTTCCGCTCGAAGTCCCAAGTGCCTTGACCTGTTTAGACACA	7696
Qy	2542	GlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnPro	2561
Db	7697	GGTGAATGGATGTAGATGCCCATCTTGCATGCCATCCACTGTCATGATCCCAACAACCC	7756
Qy	2562	IleGluAsnGlyPheValGluGlyAlaAspTrpSerTrpGlyAlaIleIleIleTrpSer	2581
Db	7757	ATTGAAATATGGTTTGTAGAGGTGCAGATTACAGCTATGGTGCATATCATCTACAGT	7816
Qy	2582	CysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrp	2601
Db	7817	TGCTTCCCTGGGTTTCAAGTGGCTGGTCATGCCATGCCAGACCTGTGAAGAGTCAGGATGG	7876
Qy	2602	SerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPhe	2621
Db	7877	TCAAGTTCCATCCCAACATGTATGCAATAGACATGTGGGCTCCCTCTCTCATATAGATTTT	7936
Qy	2622	GlyAspCysThrLysLeuLysAspAspGlnGlyTrpPheGluGlnGluAspAspMetMet	2641
Db	7937	GGAGACTGTACTAAATCTCAAGAGATGACAGGATATTTTGGAGCAAGAGAGCATGATG	7996
Qy	2642	GluValProTrpValThrProHisProProTrpHisLeuGlyAlaValAlaLysThrTrp	2661
Db	7997	GAAGTTCCATATGTGACTCTCACCTCTTATCATTTTGGAGCAGTGGCTAAACCTGG	8056
Qy	2662	GluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTrpGlyThrMetVal	2681
Db	8057	GAATAATCAAGAGAGTCTCTGCTACACATTCATCAAACTTTCTGTATGTGTACCATGGTT	8116
Qy	2682	SerTrpThrCysAsnProGlyTrpGluLeuGluAsnProValLeuIleCysGlnGlu	2701
Db	8117	TCTATACCTGTATTCAGGATATGAATCTCTGGGAGACCTGTGTCTGATCTGCCAGGA	8176
Qy	2702	AspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThr	2721
Db	8177	GATGAACCTTGGATGCGAGTCACCATCTCGCATTTCAATTTGAATGTGACTTGCCTACT	8236
Qy	2722	AlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTrp	2741
Db	8237	GCTCTCGAAAAATGGCTTTTTCGCTTTTACAGAGACTAGCATGGGAAGTGTCTGTGAGTAT	8296
Qy	2742	SerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArg	2761
Db	8297	AGCTGTAAACCTGACACATCTTAGCAGGCTCTGACTTAAGGCTTTCTCTAGAGATAGA	8356
Qy	2762	LysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnProVal	2781
Db	8357	AAGTGGAGTGGTGGCTCCCAACGCTGTGAAGCCATTTTCATGCAAAAAGCCAAATCCAGTC	8416

QY	2782	MetAsnGlySerIleIysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCys	2801	QY	3142	SerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPhe	3161
DB	8417	ATGAATGGATCCATCAAGAGGAGCAACTACATACCTCAGACAGGTGTACTATGAGTGT	8476	DB	9497	AGTGAAGTGAACCTCAGATGTCTGGAGGTTTACGATGATACAGATACAGATACATTC	9556
QY	2802	AspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrp	2821	QY	3162	ThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCys	3181
DB	8477	GACCCCGGATATGTCTGAATGGCACTGAGAGGAGAAACATGCCAGGATGACAAAACCTGG	8536	DB	9557	ACCTGTGACAGAAAGATGGTGGTCTGGTTCCTCTGAGAGAAATCTCTGTCAGTCTCTAAATAATGT	9616
QY	2822	AspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAlaAsn	2841	QY	3182	ProLeuProGluAsnIleThrHisIleLeuValHisGlyVAspAspPheSerValAsnArg	3201
DB	8537	GATGAGGATGAGCCCATTTGCCATTTCTGTGGACTGCGAGTTCACCCCGAGTCTCAGCCAAAT	8596	DB	9617	CCTCTCCGGAAACATAACACATATATCTTGTATACATGGGACGATTTCAAGTGTGAATAGG	9676
QY	2842	GlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGlu	2861	QY	3202	GlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCys	3221
DB	8597	GGCCAGGTGAGAGGACGAGTACATTCCTCAAAAAGAGATGAAATACATCTGCAATGAA	8656	DB	9677	CAAGTTTCTGTGTCATGTGCAGAGGGGTATACCTTTGAGGGAGTTTACATATCAGTATGT	9736
QY	2862	GlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGly	2881	QY	3222	GlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProValSerCys	3241
DB	8657	GGGTTCCTGCTTGAGGGAGCCAGGAGTCGGGTGTGTCTTGCCAAATGGAAGTTGGAGTGA	8716	DB	9737	CAGCTTGATGAACCTGGGAGCCACCAATCTCCGATGATCTTCGACGTCCAGTTTCTTGT	9796
QY	2882	AlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyVal	2901	QY	3242	GlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGluSer	3261
DB	8717	GCCACTCCCGACTGTCTGCTGTGCATGTGCACCCCGCCACAACTGGGCCAATGGGGTG	8776	DB	9797	GGAAAACCTGAAAGTCCGAAACATGGATTTGTGTGGCAGTAAATACACCTTTGAAAGC	9856
QY	2902	ThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyr	2921	QY	3262	ThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys	3281
DB	8777	ACGGAAGGGCTGGACTATGGCTTCATGAAGAGAGTAACATTTCCACTGTACAGAGGGCTAC	8836	DB	9857	ACAAATTATTCAGTGTGAGCTTGCTATGAACTAGAGGGGACAGGGAACTGTCTGC	9916
QY	2922	IleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIle	2941	QY	3282	GlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThr	3301
DB	8837	ANCTTGCACGGTGCTCBAAACTCACCTGTCACTCAGATGGCAACTGGGATGCAGAGATT	8896	DB	9917	CAGGAGAACAGACAGTGGAGTGGAGGGTGCATATATGCAGAGAGACAGGTGTGAAACT	9976
QY	2942	ProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsn	2961	QY	3302	ProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnVal	3321
DB	8897	CCTCTCTGTAACACAGTCAACTGTGGACCTCTCTCAAGATCTTGCCCATGGTTTCCCTAAT	8956	DB	9977	CCACTTGAATTTCTCAATGGGAAAGCTGACATTTGAAACAGGACGACTGGAGCCCAACGTG	10036
QY	2962	GlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeu	2981	QY	3322	ValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGlu	3341
DB	8957	GGTTTTTCTTTATTCATGGGGGCCATATACATACATGAGTGCATCTCTGGTTTATAAGCTC	9016	DB	10037	GTATATTCCTGCAACAGAGGCTACAGTCTTGAGGGCCATCTGAGGGCCACACTGCACAGAA	10096
QY	2982	HisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSer	3001	QY	3342	AsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPhe	3361
DB	9017	CATGGAAATTCATCAGAGAGTGCCTCTCCAAATGGCTCTCTGGAGTGGCAGTCCACCTTCC	9076	DB	10097	AATGGAACCTGGAGCCACCCAGTCCCTCTCTGCAACCAATCCATGCCCTGTCTCTTTT	10156
QY	3002	CysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAsp	3021	QY	3362	ValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAsnGlnAsnValSer	3381
DB	9077	TGCCTGCTTGCAGATGTTCCACACCAAGTAATTTGAATATGAACTGTCTCAATGGGACAGAT	9136	DB	10157	GTGATTTCCCGAGAAATGCTCTGCTGTCTGAAAGAGGATTTTATGTGATCAGAAATGTGTCC	10216
QY	3022	PheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeu	3041	QY	3382	IleLysCysArgGluGlyPheLeuGlnGlnHisGlyIleIleThrCysAsnProAsp	3401
DB	9137	TTTGACTGTGGAAAGGAGCCCGGATTCAGTGTCTCAAGGCTTCAAGCTCTCAGGACTT	9196	DB	10217	ATCAATGTAGGAGGAGGTTTCTGCTGCAGGGCCACGGCATCATTTACCTGCAACCCCGAC	10276
QY	3042	SerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHis	3061	QY	3402	GluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHis	3421
DB	9197	TCGAAATACCTGTGAGACCGATGCGCAGTGGAGCTCTGGGTTCCTCCCACTGTGAACAC	9256	DB	10277	GAGACGTGGACACAGACGCGCAATGTGAAAAAATCTCATGTGTGTCACAGCTCAC	10336
QY	3062	ThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrp	3081	QY	3422	ValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSer	3441
DB	9257	ACTTCTGTGTGTTCTCTCTCAATGATACCAAAATGCGTTTCATCAGTGAGACAGCTCTTGG	9316	DB	10337	GTAGAAATGCAATTTGCTCGAGGCGTATCATTAATATGAGACATGATCACCTACTCA	10396
QY	3082	LysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAsp	3101	QY	3442	CysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThr	3461
DB	9317	AAGGAAATGTGATATTAACCTGAGCTGCGATATGATATGATACAAAGGAGCTTCAGAT	9376	DB	10397	TGTTACAGTGGATACATGTTGGAGGTTTCTGAGGAGTGTGTTGTTAGAAATGGAACA	10456
QY	3102	LeuIleCysThrGluLysGlyValTrpSerGlnProTrpProValCysGluProLeuSer	3121	QY	3462	TrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIle	3481
DB	9377	CTGATTTGTACAGAAAGGGGTATGGAGCAGGCTTATTCAGTCTGTGAGCCCTTGTCC	9436	DB	10457	TGGACATCACTCTATTTGACAGAGCTGTCTGTCCATTTCCATGTGCAGATGGGGCATC	10516
QY	3122	CysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGlu	3141	QY	3482	CysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGluGlu	3501
DB	9437	TGTGGGTGCCACCGCTGTCTGCGCAATGTCAGTGGCACTGGAGAGGCACACACTATGAA	9496	DB	10517	TGCCACGCCCAATGCTGTCTGCTCCAGAGGGCTGATGGGGCGGCTCTCTGTGAGAGAA	10576
QY				QY	3502	ProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCysAsp	3521

```

Db 10577 CCAATCTGCATCTTCCCTGCTGTAAGAGAGTCTGTGTGGCCCTTACAGTGTGAC 10636
QY 3522 CysProGlyTThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeu 3541
Db 10637 TGCCCGCTGGCTGAGCGGGTCTGCTGTATACAGCTTTTGACGCTCTCCCTGCTTA 10696
QY 3542 AnGlyGlyLysCysValArgProHisArgCysHisCysLeuSerSerThrGlyHis 3561
Db 10697 AATGGTGAATAATGTGAAGACCAACCGATGTCACTGTCTTCTTCTTGGACGGGACAT 10756
QY 3562 AnCysSerArgLysArgThrGlyPhe 3571
Db 10757 AACTGTTCAGGAAAGGAGGACTGGGTTT 10786

RESULT 5
AAD32026
ID AAD32026 standard; cDNA; 11230 BP.
XX
AC
XX
XX
XX 18-JUN-2002 (first entry)
XX
XX Mouse C3b/C4b complement receptor-like cDNA.
XX
XX Mouse; C3b/C4b complement receptor-like protein; CR-like; diabetes;
XX immune system disorder; rheumatoid arthritis; psoriatic arthritis;
XX inflammatory arthritis; osteoarthritis; inflammatory joint disease;
XX autoimmune disease; multiple sclerosis; inflammatory bowel disease;
XX transplant rejection; graft versus host disease; atherosclerosis; lupus;
XX stroke; Alzheimer's disease; ischaemia; metabolic disorder; obesity;
XX myocardial infarction; ischaemia; metabolic disorder; obesity;
XX reproductive disorder; infertility; nervous system disorder;
XX gene therapy; immunomodulatory; antipsoriatic; antiinflammatory;
XX neuroprotective; vasodilator; cardiant; anorectic; gene; ss.
XX
XX OS Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 95..10879
XX /*tag= a
XX /product= "Mouse C3b/C4b CR-like protein"
XX /transl_except= (pos:5363..5365, aa:Xaa)
XX /note= "Xaa corresponds to an unknown amino acid; CDS
XX does not include start codon"
XX /partial
XX
XX W0200210388-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 24-JUL-2001; 2001WO-US023548.
XX
XX PR 01-AUG-2000; 2000US-0222438P.
XX
XX PA (AMGE-) AMGEN INC.
XX
XX XX Welcher AA, Elliot GS;
XX
XX DR WPI; 2002-257381/30.
XX
XX DR P-PSDB; AAE20147.
XX
XX XX Nucleic acid encoding a novel C3b/C4b Complement Receptor (CR)-like
XX nucleic acid molecule, useful for treating, preventing and diagnosing
XX rheumatoid arthritis, psoriatic arthritis, inflammatory arthritis, and
XX multiple sclerosis.
XX
XX PS Claim 1; Fig 2A-2H; 201pp; English.
XX
XX CC The invention relates to C3b/C4b complement receptor (CR)-like protein
XX and its corresponding nucleic acid sequence. The C3b/C4b CR-like
XX polypeptide and nucleic acid molecules may be used to treat, prevent,
XX ameliorate, diagnose and/or detect diseases such as immune system

```

disorders such as rheumatoid arthritis, psoriatic arthritis, inflammatory arthritis, osteoarthritis, inflammatory joint disease, autoimmune disease, multiple sclerosis, lupus, diabetes, inflammatory bowel disease, transplant rejection, graft versus host disease, nervous system disorders (e.g. stroke, Alzheimer's disease), ischaemic conditions (e.g. atherosclerosis, stenosis, myocardial infarction, and ischaemia), metabolic disorders (e.g. obesity and diabetes); and reproductive disorders and infertility. The C3b/C4b CR-like nucleic acid molecules are used in gene therapy. The present sequence is mouse C3b/C4b CR-like cDNA

XX Sequence 11230 BP; 2714 A; 2981 C; 2886 G; 2648 T; 0 U; 1 Other;

Alignment Scores:

Score: 0 Length: 11230
 Pred. No.: 16707.00 Matches: 2915
 Percent Similarity: 89.72% Conservatives: 289
 Best Local Similarity: 81.63% Mismatches: 361
 Query Match: 83.65% Indels: 6
 DB: 6 Gaps: 5

US-09-977-053-4 (1-3571) x AAD32026 (1-11230)

```

QY 1 MetTyrProArgLeuAlaPheCysCysTyrGlyLeuAlaLeuValSerGlyTrrpalaThr 20
Db 176 ATGTGTGGCTGGCTGGCTTTGTGTGGCTCTGGCACTGTGTGGCTGGCAAC 235
QY 21 PheGlnGlnMetSerProSerArgPheSerPheArgLeuPheProGluThrAlaPro 40
Db 236 TTCCAGCCGCTGGCCCTTGGCTCAACTTCAGCTTCGGCTTCGGCTTCGGCGGCTCTCG 295
QY 41 GlyAlaProGlySerIleProAlaProAlaProGly--AspGluAlaAlaGlySer 59
Db 296 GGGGCTCTGGGCGAGCTGGCGGTACTCTCCGGCTCCAGTGGAGGAGGAGGAGGAGGAGC 355
QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgValArgLeuArgGluLeuSer 79
Db 356 AAAAGTGGAGGCTGGCGCGCGTTCGGAGCCGGTTCGGAGCGCGTGGCGGAGCTCAGC 415
QY 80 GluArgLeuGluLeuValPheLeuValAspSerSerSerValGlyGluValAsnPhe 99
Db 416 GGCAGCTGGAGCTGTCTCTCTGGTGGAGAGTGTCTCCAGCGTGGCGCCAAACCACTTC 475
QY 100 ArgSerGluLeuMetPheValArgValLeuLeuSerAspPheProValProThrAla 119
Db 476 CTCACAGAGCTCAAGTGTGTGGCAAGCTGTCTCCGACTTCCCTGGTGTGTCCAGCGCC 535
QY 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139
Db 536 AGCGTGTGGCCATCTGTCACTCTCATCAAGAACCAACGTTGGTGGCGCGGTGGATTAC 595
QY 140 IleSerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAla 159
Db 596 ATCTCCACAGCGCGCGCCACCAACACAGTGTGGCTGTCTCAGCCCGAGATCCCGGCC 655
QY 160 IleSerTyrArgGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
Db 656 ATCACTTACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 715
QY 180 LeuLeuHisAlaArgGluAsnSerThrLysValValPheLeuLeuThrAspGlyTyrSer 199
Db 716 CTTCGTCACTCTAGAGAAACTCCACCAAGTCAATATTTCTCATCCAGCGCTATTATCC 775
QY 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
Db 776 AATGGCGAGAGCCCGAGACCTATTTCAGCAGCAGCTTCGGGATTTCCGAGTGGAGATCTTC 835
QY 220 ThrPheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
Db 836 AGTTTGGGATTTGGCAGGGGATATTCGGAACTGATGATGATGATGATGATGATGATGAT 895
QY 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAla 259
Db 896 GAAGAACAATTGTACTGTCTCCACAGTTTGAAGAAATTTAGAGCTTTAGCTTCGAGGGCG 955

```

QY 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
DB 956 TTGCATGAAGATCTACCTCTCGGAGTTTATCCAGAGATATGGCCCACTGCTCTTAT 1015
QY 280 LeuCysAspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHis 299
DB 1016 CTCTGTGAGGCTGGGAAGAGCTCTGTGACAGAAATGGCCAGCTGCAAAATGTGGACACAC 1075
QY 300 ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGlu 319
DB 1076 ACGGGTCAATTTGAAATGTCATCTGTGGAAGGGCTATTACGGGAAGGCTGCGACATGAG 1135
QY 320 CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer 339
DB 1136 TGCACAGCTTGGCCATCAGGACATATACCGGAGCTTCTCAGAGGAATCAGCC 1195
QY 340 CysIleProCysProAspGluAenHisThrSerProGlySerThrSerProGluAsp 359
DB 1196 TGCATCCATGCTCTGAGGTAAAGCCACACCTCCCACTTGGAGCACTTCCCTGGAAGAC 1255
QY 360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
DB 1256 TGGGTGTGGCCAGAGGATACAGAGATCTGGCCAGACCTGTGAGGTGTGCCACTGTCT 1315
QY 380 AlaLeuLysProGluAenGlyTyrPheIleGlnAenThrCysAenAenHisPheAen 399
DB 1316 GCCCTGAGGCTCTGAAATATGTTTTTTTATACAAACACTTGCAGAAACTTACTTCAT 1375
QY 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu 419
DB 1376 GCGCCTGTGGGTCCGATGTCGCCCGGGCTTGGACCTTGTGGAGCAGCATCCATTG 1435
QY 420 CysLeuProAenGlyLeuTrpSerGlySerGlySerTyrCysArgValArgThrCysPro 439
DB 1436 TGTCAACCCCAATGGTTGTGTGGTGTGGACAGAAAGCTTCTGCAGAGTGAAGAGTGGCCC 1495
QY 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys 459
DB 1496 CACCTCCGACAGCCCAACACGCGCCACATCAGCTGTCTCCACTCGGGAATGTCTCTACAC 1555
QY 460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr 479
DB 1556 ACCCTGTGTTGTTGTTACCTGCAATGAAGATACAGATTAGAGGCAGCAGCTAGGCTTACC 1615
QY 480 CysGlnGlyAenSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSer 499
DB 1616 TGTCAAGGAAATGCCAGTGGGATGGCCAGAGCCCGGTGTGTAGAACGCCATTTGTGCC 1675
QY 500 ThrPheGlnMetProLysAspValIleIleSerProHisAenCysGlyLysGlnProAla 519
DB 1676 ACCTCCAGAGCCCAAGCGGTCATCATTTCTCCACCCAGCTGGCGAGCAGCCGCC 1735
QY 520 LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
DB 1736 AGGCTGGGATGACCTGTACAGTAAGCTGCGCGCCAGGATACATTTTATCGGGGTGAGA 1795
QY 540 GluMetLeuArgCysThrThrSerGlyLysThrPheAsnValGlyValGlnAlaValCys 559
DB 1796 GAA--GTGAGATGTGCCATCTCGGGAAGTGGAGTGCCAAAGTTTCAGACAGCTGTGTGC 1852
QY 560 LysAspValGluAlaProGlnIleAenCysProLysAspIleGluAlaLysThrLeuGlu 579
DB 1853 AAGATGTGGAGCTCCACAAATCAGCTGTCTCAATGACATGAGGCAAGATCTGGGAG 1912
QY 580 GlnGlnAspSerAlaAenValThrTrpGlnIleProThrAlaLysAspAenSerGlyGlu 599
DB 1913 CAGCAGGACTCTGCTAATGCCACTGCGAAGTCCCAACAGCTAAAGACAACTCTGGTGA 1972
QY 600 LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
DB 1973 AAGGTGTCTAGTCCAGCTCCACCCAGCTTTACCCCACTTACCTCTTCCCAATTTGGAGAC 2032

QY 620 ValAlaIleValTyrThrAlaThrAspLeuSerGlyAenGlnAlaSerCysIlePheHis 639
DB 2033 GTGGCCATCACCTACAGCGCAACCGACTCATCCGTTAACCAAGCCAGCTGCATTTCTAC 2092
QY 640 IleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProProVal 659
DB 2093 ATTAAAGTCAATTGATGTGGAACCGCTGTATAGATTGTGTCGCGATCTCCACCTCAATC 2152
QY 660 GlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAenSer 679
DB 2153 CAGTCTGAGAGAGGAGCCCTTGCAGCTGGGATGAGCTCAGTCTTCAGACAACTCC 2212
QY 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
DB 2213 GGGGCTGAATTTGCTCAATTACAGCAGTACACACAGCGGACATGTTTCTCATGGGAA 2272
QY 700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAenAenArgThrCysAspIleHis 719
DB 2273 ACGTGTGTGTGTACACAGCCACTGACCCCTCAGGCAACCAACAGGACCTGTGACATCCAC 2332
QY 720 IleValIleLysGlySerProCysGluIleProPheThrProValAenGlyAspPheIle 739
DB 2333 ATTGTCAATAAAGGTTCTCCCTGTGAGGTCCTTCAACCTGTAAACGGGAGCTTTATC 2392
QY 740 CysThrProAenThrGlyValAenCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
DB 2393 TGTGCCCAGATAGTGTGGAGTTAACTGTAGCTGAGCTGCAAGAGGGCTATGATTTC 2452
QY 760 ThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThr 779
DB 2453 ACAGAGGGTCACTCAGAAAGTACTACTGTGCTTTTGAAGATGTGTATCTGGAGACCA 2512
QY 780 TyrThrThrGluTrpProAspCysAlaLysLysArgPheAlaAenHisGlyPheLysSer 799
DB 2513 TACTTACAGATGGCCAGACTGTGTCTATAAAGCTTTTGCACACCATGGTTTCAAGTCC 2572
QY 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
DB 2573 TTTGAAATGCTATACAAACCACTCGCTGTGTGATGACATGATCTGTTTAAAGATTTTCT 2632
QY 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
DB 2633 GCAGCATTTTCAGACTACCTCGGGGAACATGTCCTCCGCTCTTTGTAACGATGCTGATGAC 2692
QY 840 IleAspCysArgLeuGluAenLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
DB 2693 ATTGACTGACAGCTGGAGGAC---CTGACCAAAAAATACTGATCGAGTATATATTACAA 2749
QY 860 TyrGluAenGlyPheAlaIleGlyProGlyTyrTrpGlyAlaAlaAenArgLeuAspTyr 879
DB 2750 TATGAAATGGCTTTGCAATTGGACAGGAGCTGGGGTGCAGGCAACAGGCTGGATTAT 2809
QY 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
DB 2810 TCCTACGATCACTTCTCGATGTTGTACAGGAACACCCAGCATGTGGGCAAGGCCAGA 2869
QY 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAen 919
DB 2870 TCTGTACGGAATTAAGAACTGTCCCATTTGTGACCCCAAAATTCAGTAATTTTAAAC 2929
QY 920 IleThrAlaSerValProLeuProAspGluArgAenAspThrLeuGluTrpGluAenGln 939
DB 2930 ATCAGCTAGCTGTGCCACTCCAGAGGAAGAAACGATACCTTTGAATTGGAGAAATCAG 2989
QY 940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLysLysArgThrLeuAenLys 959
DB 2990 CAGCGACTCATTAAGACATTTGGAACAATCAACCAATCGCTGGAAGAACCCCTTGAATAA 3049
QY 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
DB 3050 GAGCCCATGATTTCTTTCCAGCTCGCTCGAACAACAGTGTGGCTGACAGCAATTCCTC 3109
QY 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999

Db 3110 GAAACAGAAAAGGCTTTCTCTCTGACAGCAGGCTCTGCTGAGGGGGCGCATGTGT 3169
Qy ValAenCysProLeuGlyThrTyrrAsnLeuGluHisPheThrCysGluSerCysArg 1019
Db 3170 GTCAATGCGCCCTGGGAACCTCTTACTCTCTGGAGCATTCACCTGTGAAAGCTGCCTC 3229
Qy 1020 IleGlySerTyrrGluAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMet 1039
Db 3230 ATGGGATCTACCAAGATGAAGAGGGCAGCTGGAATGCAAGCTCTGTCCCCAAGGACT 3289
Qy 1040 TyrThrGluTyrrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
Db 3290 CACGGCGAATACCTCCATTCAGAAAGCGCTCTCTGAATGCAAGCTCAGTGTAAAGCAAGGC 3349
Qy 1060 ThrTyrrSerTyrrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrrGlnPro 1079
Db 3350 ACCTACTCTTCAGTGGGCTGGAGACCTGCGAATGTGTGCGCTGGTACTTATCAACGG 3409
Qy 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
Db 3410 GAAATTGGATCCCGAGCTGCTCTATGCCAGAAAACCCACCAACGGTGAAGAAGGA 3469
Qy 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
Db 3470 GCGGTGGACATCTCTGCTTGTGGAGTGGCTGCCAGTAGAGAAATCTCCCGTTCTGGG 3529
Qy 1120 LeuMetProCysHisProCysProArgAspTyrrGlnProAsnAlaGlyLysAlaPhe 1139
Db 3530 CTAAACACCCCTGCTACCTTGGCCCTCGAGACTATTACCAACCCAGTGGAGGAAAGCTTTC 3589
Qy 1140 CysLeuAlaCysProPheTyrrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
Db 3590 TGCCCTGGCTTGTCCCTTTATGGAATCAACCATCACTGCGGCCCAAGTCCATCACAGAC 3649
Qy 1160 CysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValProProAlaSer 1179
Db 3650 TGCTCAAGTTTATGCTTACTTCTCAGCAGCAGAAAGAAAGCATAGTGCCTCGTGCC 3709
Qy 1180 LeuGlyHisIleLysLysArgHisGluIleSerGlnValPheHisGluCysPhePhe 1199
Db 3710 CTGGACATTCGCAAGAACGATACGAGTACGAGTACAGTCTTTTACGAAATGCTCTTCA 3769
Qy 1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrrValCysLeuCys 1219
Db 3770 AACCCCTGCCAACACAGTGGACCTGCCAACAGCTTGGCGGTGGTATGCTCTCTGTC 3829
Qy 1220 ProLeuGlyTyrrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
Db 3830 CCACCTGGATACACAGGCTTAAAGTGTGAACAGATATTGATGAATGCAAGCTCTCTGCT 3889
Qy 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
Db 3890 TGCCCTCAATGGTGGAAATTTAGAGACCCAGTTGGGGGATTCACGTGCGAATTTCAATG 3949
Qy 1260 GlyTyrrThrGlyLysArgCysGluGluAsnIleAsnGluCysSerSerProCysLeu 1279
Db 3950 GGTATTTCAGGTCAAAATATGTGAAGAAATATAAATAGTGTATCTCCAGCCCTTGTCTTA 4009
Qy 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrrArgCysThrCysValLysGlyPhe 1299
Db 4010 AATAAAGGAACCTGCACCTGACGGCTTGGCAAGCTACCCGCTGTACCTGTGTGAAGAGATAC 4069
Qy 1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
Db 4070 ATGGGTGTGCATCTGTGAACACAGCTCATGATGATGCCAGTCAAGCCCTCTTAAACAC 4129
Qy 1320 AlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGly 1339
Db 4130 GCAGTTGTGAAGACCAAGCTGGGGGGTCTCGTGCATAATGCCAACCCCGGATTTTGGGT 4189
Qy 1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr 1359

Db 4190 ACTCGGTGTGAAAAAATGTGGATGAGTGTCTCAGTCAGCCATGCCAAAATGGAGCCACT 4249
Qy 1360 CysLysAspGlyValAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
Db 4250 TGTAAAGATGGTGGCCACACAGCTTCAGGTGTCAATGTCCAGCAGGCTTCACAGGGACAC 4309
Qy 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
Db 4310 TGTGAACCTGAACATCAACGAGTGTCACTCAACCCGCTAGGAAACAGGCCACCTGTGTG 4369
Qy 1400 AspGluLeuAsnSerTyrrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db 4370 GATGAACCTAAACTCATACAGTTGTAAATGTACGCCAGGATTTTCAGGCCACAGGTGTGAG 4429
Qy 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrrGlyTyrrVal 1439
Db 4430 ACAGAACAGCTTCGCGTTTAACTGTGATTTTGAAGTTTCTGGCATCTACGGGTACGTC 4489
Qy 1440 MetLeuAspGlyMetLeuProSerIleHisAlaLeuThrCysThrPheThrMetLysSer 1459
Db 4490 CTGCTAGATGGAGTGTGCCAACCTCCATGCCATGCCATAACCTGGCATTTCTGGATGAATCC 4549
Qy 1460 SerAspAspMetAsnTyrrGlyThrProIleSerTyrrAlaValAspAsnGlySerAspAsn 1479
Db 4550 TCTGATGTCACTCAACTACGGGACGCCCATCTCTATGCACTTGAGGATGACAAAGACAAC 4609
Qy 1480 ThrLeuLeuLeuThrAspTyrrAsnGlyTyrrValLeuTyrrValAsnGlyArgGluLysIle 1499
Db 4610 ACCTTCTCTCTGACTGATTAACAAGCTGGGTCTTTATGTGAATCGAAAGAAAGATC 4669
Qy 1500 ThrAsnCysProSerValAsnAspGlyArgTyrrHisIleAlaIleThrTyrrThrSer 1519
Db 4670 ACCAACTGCCCTCCGTTAAATGATGTCATTTGGCATCATATTGCAATCATCATGACAGT 4729
Qy 1520 AlaAsnGlyIleTyrrLysValTyrrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeu 1539
Db 4730 ATTGTGGAGCTTGGAGGTCTATATAGATGGGAAATATCTGACCGTGTGTACTGCGCTC 4789
Qy 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLys 1559
Db 4790 TCCATTGGCAAGCCATCTGCTGGCGTGCATTAGTTCTTGGGCAAGCAAGCAAA 4849
Qy 1560 LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579
Db 4850 AAAGGAGAGGGGTTCACCCGGCTGAGTCTTTGTGGGGCTCCATATAGCCAGCTCAACCTC 4909
Qy 1580 TrpAspTyrrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4910 TGGGACTATGCTCTCTCTCCACAGCAGGTGAAGTGTCTGCGCCAGCTCTCTGCCAGAGAA 4969
Qy 1600 LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysVal 1619
Db 4970 CTGAGTCGGGAAAACGTGTAGCATGGCCGATTTCTGTCGGGAATCAGCGGGAAGGTG 5029
Qy 1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro 1639
Db 5030 AAGGTGTATTCAGCAGCATGTCTGCTGATGTCCCTTTAGAGATCCGTCCT 5089
Qy 1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
Db 5090 CACTTGAGACCTGTCATCAGAAATCGMAAGCCAGGCTCCAAAGTCAGTCTGTTCTGTAT 5149
Qy 1660 ProGlyPheGlnLeuValGlyAsnProValGlnTyrrCysLeuAsnGlnGlyGlnTyrrThr 1679
Db 5150 CCGGGCTTCAGATGTGGGAATCTGTGTCAGTATTTGTCTGAACCAAGGGCGATGGACA 5209
Qy 1680 GlnProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGly 1699
Db 5210 CAACCACTCCCTCATGTGAACGCAATTCCTGTGGCTGCCCTCCGCCCTTGGAGATGGC 5269
Qy 1700 PheHisSerAlaAspPheTyrrAlaGlySerThrValThrTyrrGlnCysAsnAsnGly 1719
Db 5270 TTCTACTCAGCGAGGACTTCCATGGGGCAGCAGCGTGCACCTATCAGTGCACCACTGGC 5329

Db 7490 TGTCCCCAACCTGAGGAGATCTCAACGGTATCATCCACGACACAGGGCTTGCCCTATCTC 7549
Qy SerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeu 2479
Db 7550 AGCACCACTCTACACCTGACAGCCAGGCTTGAGTTAGTGGGCAATGCTACCAACCCCTC 7609
Qy CysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeu 2499
Db 7610 TGTGGGAAATGGCCAGTGGCTCGAGAGAAACCAATGTGCAAAACCAATGGAATGGCCCA 7669
Qy LysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr 2519
Db 7670 GAGCCCAAGGAGATTTAAATGGCCCAATCTCTCCGTGAGCTTCAGTATGGAACAACC 7729
Qy ValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
Db 7730 ATCACAATCTTTGTGACCGGGCTTCGGCTCGAAGGTCCCAATCCCTGACCTGTGTTA 7789
Qy GluThrGlyAspTyrAspValAspAlaProSerCysAsnAlaIleHisCysAspSerPro 2559
Db 7790 GAGACAGGTGATGGGATATGATCCCTCTTGTGATGCCATCCACTGCAGTGACCCA 7849
Qy GlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIle 2579
Db 7850 CAGCCCATTTGAAATGGTTTGTAGAGGTGGGATTAAGATACGGTGGCCATGATCATC 7909
Qy TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
Db 7910 TATAGCTCTTCCCTGGGTTTCAGGTGCTTGGTCATGCCATGCACACCTGTGGAAGGTGG 7969
Qy GlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIle 2619
Db 7970 GGATGGTCAAGCTCCAGCCCAACCTGTGTACCATAGACTGGGGTCTCCCTCTCCACATA 8029
Qy AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAsp 2639
Db 8030 GACTTGTGTGACTGACTTAAGTCAAGATGAGTGGCAGGAGACATTTGATCAAGAGTGCAC 8089
Qy MetMetGluValProTyrValThrProHisProTyrHisLeuGlyAlaValAlaLys 2659
Db 8090 ATGATGGAAGTCCCATATCTGGCT-----CACCCCTCAACATTTGGAGCAACAGCTAAG 8143
Qy ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThr 2679
Db 8144 GCCTTGGAAATATACAAAGGAGTGGCTGCTGCATGCATGCCATCCACTCTCTATGGCAGC 8203
Qy MetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCys 2699
Db 8204 ATGGTTTCTCAGCTGGAGCTGCTGTTATGAACTGCTGGGAATCCCTGTGCTGATCTGC 8263
Qy GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
Db 8264 CAGGAAGATGTTACCTGGAAATGGTACCGCACCTCTCTGCAATTTCCATTTGAAATGATTTG 8323
Qy ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
Db 8324 CCTGTGCTCCGAAATGGCTTTTATCATTTTCACAGACGACATATGGGAGTGGCTGCA 8383
Qy GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
Db 8384 CAATATAGCTGCAAGCGGGGACATTTCTAGAAAGGCTCCCACTTAAGACTCTGTCTGCAG 8443
Qy AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
Db 8444 AATAAGCAGTGGAGTGGCACTGTTCCACGCTGTGAAGCCATCTCATGCAAGTAAAGCAAC 8503
Qy ProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyr 2799
Db 8504 CCACCTCGGAATGGATCCAAAGAGATGACTACTCTCTACTCTGGGTGTGTATCTAC 8563
Qy GluCysAspProGlyTyrValLeuAsnGlyThrGluArgThrCysGlnAspAspLys 2819
Db 8563 ACATTCACCTGCCAGCAAGATGGCCATTTGGGTCCCTCGAAAGAAATCACCTGCAGTCTCTAA 9700

Db 8564 GAGTGTGACTCTGGCTATATTCTCAATGGCTCTAAGAAGAGGACATGCCAAGAAATAGA 8623
Qy AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
Db 8624 GATTGGATGGCATAGAGCCCATATGTTATCTCTGTAGACTGTGGCTCAACCCCACTGCC 8683
Qy AlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
Db 8684 ACCAATGGCCAGTGAAGGAGAGATACATTTCCAAAGAGATTTACATCTCTTCC 8743
Qy AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
Db 8744 CGTGAAGGTTCACTTGGAGAGGACGAGGAGTGTATCTGTCTTACCAATGGAAGTTGG 8803
Qy SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsn 2899
Db 8804 AGTGTGCACTCCCGAGCTGATGCTGTAGATGCTCTGCTCCGCCACCAAGGTGCAAT 8863
Qy GlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu 2919
Db 8864 GGGTGGCAGATGGCTTAGACTATGGGTTCAAGAGAGAGTAGCGTTCCACTGTCTAGAG 8923
Qy GlyTyrIleLeuHisGlyValaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
Db 8924 GCCTATGTGCTGCAGGGGGCTCCCAAGACTCACTGTCTAGTCCAAATGGGACTTGGGATGA 8983
Qy GluIleProLeuCysLysProValAsnCysGlyProGluAspLeuAlaHisGlyPhe 2959
Db 8984 GAAGTCCCTGTCTTAAACACGACTTACCTGTGTGCTCTCTCCGCGACCTTCCCGAGGCTTC 9043
Qy ProAsnGlyPheSerPheIleHisGlyLysIleGlnTyrGlnCysPheProGlyTyr 2979
Db 9044 CCTATGTGCTTTCTTTTATATGCGGGGCCACATACAGTATCAGTGTTTTACTGGTTAT 9103
Qy LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
Db 9104 AAGCTTTCATGNAACCCCATCAAGAGATGCTTCCCAATGGTTCCTCGAGCGCAGCTCG 9163
Qy ProSerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGly 3019
Db 9164 CCATCTGCTCCCTACCTGTGAGGTGTTCACACCCATCATTTCAACAGGGACCAACACGCA 9223
Qy ThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeu 3039
Db 9224 ACTGATTTGGGATGTGGAAGACGCTCCAGATTGAGTGTCTTCAAGGCTTCAAGCTGCTT 9283
Qy GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCys 3059
Db 9284 GCACTTCTGAAATCACCTGTGATGCAATGCAATGGTCTGAC---GTCCCACTGTGT 9340
Qy GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer 3079
Db 9341 GAGCACGCTCAGTGGGGCTCTCCCAACCATACCAACGCAATTTGCTCTGAGGGCAGC 9400
Qy SerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySer 3099
Db 9401 CTTTCCGAGGACATGTGTAACTTACAGTGCAGACCTGGCTACACCATCAAGGTAGT 9460
Qy SerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluPro 3119
Db 9461 TCAGATCTGATTTGTACGNAAGCATATGAGCCAGCCTTACCCACAGTGTGAAACC 9520
Qy LeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThr 3139
Db 9521 CTGTCTGTGGACCCCAACCACTAGCCATGAGTGGGCAACAGGAGGCTCATACC 9580
Qy TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp 3159
Db 9581 TATGAAGCAAGTCAAACTCAGGTGTCTGGAAGGTATGTGATGATTCGATACAGAT 9640
Qy ThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLys 3179
Db 9641 ACATTCACCTGCCAGCAAGATGGCCATTTGGGTCCCTCGAAAGAAATCACCTGCAGTCTCTAA 9700

QY 3180 LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerVal 3199
 Db 9701 AATATCCCTCGCATCCACATGACACCGATAGCTTTTTCAGGAGATGACTTCCAGGTG 9760
 QY 3200 AsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSer 3219
 Db 9761 AACAGACAAGTTCTGTGTATGTGTCAGAGAGGGTTTATCCCAAGAGAGTGAAGTGTCA 9820
 QY 3220 ValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProVal 3239
 Db 9821 ACATGCCAGCCCGACGATACATGAGCCACCATTTTCTGATGAATCCTGTATCCCAATT 9880
 QY 3240 SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPhe 3259
 Db 9881 GTTTGTGGGCATCTCTGAAGCCACGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9940
 QY 3260 GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg 3279
 Db 9941 GGAAGCACCATTTTATACCATGTGTACCTCTGCTTACAAATTTAGAGGGGAACAGGACGA 10000
 QY 3280 ValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLysGluThrArgCys 3299
 Db 10001 ATCTGCCAGGAGAACACAGCATGTGAGTGGAGAGTGGCAGTGTGCAGAGAGAACAGATGT 10060
 QY 3300 GluThrProLeuGluPheLeuAsnGlyValAspIleGluAsnArgThrThrGlyPro 3319
 Db 10061 GAGACTCCAGCTGAGTTTCCATGGAAGGCTGTCTTGAAGAACACACCATCTGGAGACC 10120
 QY 3320 AsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCys 3339
 Db 10121 AGCCTTCTGTTTCTCTGTCAGAGGCTACACCTCTGAGAGGTCCTCCGAGGACACATGC 10180
 QY 3340 ThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProVal 3359
 Db 10181 ACTGCAAAATGGAACCTGGAATCACTGACTCCCTCTGCAACCAATCCATGCTGCTGTC 10240
 QY 3360 ProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsn 3379
 Db 10241 CCTTTGTGATCTCTGAGAACCGCTCTTCTGAAAGAGTTTATGTCGACCAAGAT 10300
 QY 3380 ValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsn 3399
 Db 10301 GTATCTATCAAGTGCAGGAGGCTTCTGCTCAAGGCAATGCTGTCATCAGTGCAGC 10360
 QY 3400 ProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPro 3419
 Db 10361 CCTGACGAGACATGGAGCCACCAATGCCAGATGTGAAAAAATCTCTGTGTCTCTCCA 10420
 QY 3420 AlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThr 3439
 Db 10421 AGTCACGTAGAGAAATGCAATGTCTGAGGAGTGTATTACCAAGTATGGGACATGATCACC 10480
 QY 3440 TyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsn 3459
 Db 10481 TACTCTGTACAGTGTACATGCTAGAGGTTCCCTCCGAGTGTTCCTAGAGAAAT 10540
 QY 3460 GlyThrTrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGly 3479
 Db 10541 GGAACATGACACCATCTCTCTGTGTCAGAGCTGTCTGTGCTGCTGCTGCTGCTGCTGCTG 10600
 QY 3480 GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCys 3499
 Db 10601 GGTGTCTGTCAAGCTCAAAAGCTGTCTATGCCAGAGCGCTGATGGAGCGTCTCTGT 10660
 QY 3500 GluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGln 3519
 Db 10661 GAAGAGCCATATGCATACTCTCTGTTGAATGTTGGGCGCTGTGTGGCCCTTATCAG 10720
 QY 3520 CysAspCysProProGlyThrThrGlySerArgCysHisThrAlaValCysGlnSerPro 3539
 Db 10721 TGTGACTGCCCCCAAGCTGAGCTGGGTGCCGCTGTCTATCAGCTACTTGTGCTGCTGCTGCTG 10780

QY 3540 CysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThr 3559
 Db 10781 TGCTTAATGCGGGGAATGCAATAGACCAACCAACGATGCTTCTCTGACCTGGACA 10840
 QY 3560 GlyHisAsnCysSerArgLysArgArgThrGly 3570
 Db 10841 GGACATGATGTTCCAGGAAAGAGAGCGCGG 10873
 RESULT 6
 ID AAK94920 standard; cDNA; 5421 BP.
 AC AAK94920;
 XX 06-NOV-2001 (first entry)
 DT Human full-length cDNA, SEQ ID NO: 4152.
 XX Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 OS Homo sapiens.
 XX EF1130094-A2.
 XX 05-SEP-2001.
 XX 07-JUL-2000; 2000EP-00114089.
 XX 08-JUL-1999; 99JP-00194486.
 PR 11-JAN-2000; 2000JP-00118774.
 PR 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI: 2001-524255/58.
 DR P-PSDB; AAM93954.
 PT 830 Primers useful for synthesizing full length cDNA clones and their use
 in genetic manipulation.
 XX Claim 8; SEQ ID NO 4152; 1380pp + Sequence Listing; English.
 CC The invention relates to primers for synthesizing full length cDNA
 clones. 830 cDNA molecules encoding a human protein have been isolated
 and nucleotide sequences of 5' - and 3' - ends of the cDNA molecules have
 been determined. Primers for synthesizing the full length cDNA are useful
 for clarifying the function of the protein encoded by the cDNA. The full
 length clones were obtained by construction of full length enriched cDNA
 libraries that were synthesised by the oligo-capping method. The primers
 enable the production of the full length cDNA easily without any special
 methods. The present sequence is a full length human cDNA of the
 invention. Note: The sequence data for this patent did not form part of
 the printed specification, but was obtained in CD-ROM format directly
 from EPO
 XX SQ Sequence 5421 BP; 1398 A; 1275 C; 1318 G; 1430 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 5421
 Score: 9400.00 Matches: 1676
 Percent Similarity: 97.56% Conservative: 4
 Best Local Similarity: 97.33% Mismatches: 11
 Query Match: 47.06% Indels: 32
 DB: 4 Gaps: 1
 US-09-977-053-4 (1-3571) x AAK94920 (1-5421)
 QY 1881 GlyAspLysGluSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCys 1900
 Db 3 GGTGATAAAGAAATCATCTCTGCTTGTGCTTGTGAGTTCATTCCTCCCTCTGTGTGT 62

QY 1901 GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly 1920
DB 63 GAACCAAGTCAAGTCTTAGTCGGGAAATATAAATAATGGAATAATATATTTAGTGGG 122
QY 1921 LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSer----- 1936
DB 123 CTATACCTCTTCTACTGCATCATATTCATGCGATACAGGATACAG-ATGCGGGTTTCG 181
QY 1936 ----- 1936
DB 182 CCATTTTGTCCAGACTAGTCTCGAACTCTTGAGCTCAAGCAATCCACCTGCCTCAACCTC 241
QY 1937 -----LeuGlnGlyProSerIleIleGluCysThrAlaSerGly 1949
DB 242 CCAAAAGTCTGGAAATTACAGCTTACAGGGCCCTTCCATTAATTGAATGACAGGCTTCTGGC 301
QY 1950 IleTyrAspArgAlaProProAlaCysHleLeuValPheCysGlyGluProProAlaIle 1969
DB 302 ATCTGGGACAAAGCGCCACCTGCTGTCACTCTGTCTGAGGAAACACCTGCCATC 361
QY 1970 LysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsnThrValThrTyrThrCys 1989
DB 362 AAAGATGCTGTCTATTACGGGGATAAATCACTTTTCAGGACACCGTCACTTACACTTGC 421
QY 1990 LysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTyr 2009
DB 422 AAAGAAGGCTATACTTCTGCTGTGTGACCACTTGAATGCTGCGCCAGCGCAAGTGG 481
QY 2010 SerArgSerAspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHis 2029
DB 482 AGTAGAAGTGACAGCAGTGCCTGCGCTGTCTCTGTGATGAGCCACCACTTGGACCAAC 541
QY 2030 AlaSerProGluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAsp 2049
DB 542 GCCTCTCCAGAGTCTGCCATCGGCTCTTTGGAGACATTGCATTTCTACTGCTCTGAT 601
QY 2050 GlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTyrValPro 2069
DB 602 GGTACAGCTTACAGACAAATCCAGCTTCTCTGCAATGCCAGGCGCAAGTGGGTACCC 661
QY 2070 ProGluGlyGlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerVal 2089
DB 662 CCAGAAGGTACAGACAAATCCAGCTTCTCTGCAATGCCAGGCGCAAGTGGGTACCC 721
QY 2090 SerTyrSerIleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSer 2109
DB 722 TCCTATAGCATCTTGGAACTCTGTAGCAAGCAAAATTTGCAGCTGGCTCAGTTGTGAGC 781
QY 2110 PheLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGly 2129
DB 782 TTTAAATGCAATGGAGGCTTTGTACTGAAACACCTCAGAGTCAATGAATGTATGAGAGGT 841
QY 2130 GlyGlnTyrAsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluPro 2149
DB 842 GGGCAGTGGAAACCTTCCCGATGTCCATCCAGTGCATCCCTGTGGGGTGTGGAGGCCA 901
QY 2150 ProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAla 2169
DB 902 CCAAGCATCATGAATGGCTATGCAAGTGGATCAAACTACAGTTTTTGGAGCCATGGTGGCT 961
QY 2170 TyrSerCysAsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThr 2189
DB 962 TACAGCTGCAACAGGGGTTTACATCAAAAGGGGAAAGAGAGCACCCTGGCAGGCCCTCA 1021
QY 2190 GlyGlnTyrSerProIleProThrCysHisProValSerCysGlyGluProProLys 2209
DB 1022 GGGCAGTGGAGTGTCTATACGAGCTGCCACCCGGTATCTTGTGGTGGACCACTTAG 1081
QY 2210 ValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyr 2229
DB 1082 GTTGAGATGGCTTTCTGAGCAGATCAAACTGGCAGGATCTTTGAGAGTGAAGTGGATAT 1141

QY 2230 GlnCysAsnProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArg 2249
DB 1142 CAGTGTAAACCCGGGCTATTAAGCCAGTCGGAAGTCTGTATTTGTCTGCCATGCCAATGCG 1201
QY 2250 HisTyrHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysProProPro 2269
DB 1202 CACTGGCACAAGTGAATCCCTCTGATGTGGGTCTCTCTCGACTGTGGAAAACCTCCCGG 1261
QY 2270 IleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePhe 2289
DB 1262 ATCCAGATGGCTTCAATGAAGGAGAAATCTTGAAGTAGGGTCCAAAGGTTCAAGTTTTC 1321
QY 2290 CysAsnGluGlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLys 2309
DB 1322 TGTAAATGAGGGTTATGAGCTTGTGTGACAGTCTTGTGACATGTTCAGAAATCTGGCAAA 1381
QY 2310 TrpAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeu 2329
DB 1382 TGGAAATAGAGTCAAAATCCAAAGTGCATGCTGCCAAGTCCCGAGGCCGCCCTCTTG 1441
QY 2330 GluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCys 2349
DB 1442 GAAAACCACTAGTATTAAGGAGTTCACCAACGAGTAGGAGTTGTGACATTTTCCTGT 1501
QY 2350 LysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyr 2369
DB 1502 AAAGAAGGCGGTGCTGTCAGAGCCCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1561
QY 2370 AsnAspSerPheProValCysLysIleValLeuCysThrProProProLeuLysSerPhe 2389
DB 1562 AATGACTCTTCCCTGTTTGAAGATTGTTCTTTGTACCCCACTCCCTTAATTTCTCTT 1621
QY 2390 GlyValProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysVal 2409
DB 1622 GGTGTCCCATCTCTTCTGCTCTCTCAATTTTGGAAAGTACTGTCAAGTATTTCTGTGTA 1681
QY 2410 GlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTyrSer 2429
DB 1682 GGTGGGTTTTCTTAGAGGAAATCTACACCTCTGCGACCTGTATGACACCTGAGC 1741
QY 2430 SerProLeuProGluCysValProValGluCysProGlnProGluLysLeuProAsnGly 2449
DB 1742 TCTCCACTCCAGAAATGCTTCCAGTAGAATGTCCCAACCTCAGGAAATCCCAATGGA 1801
QY 2450 IleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGly 2469
DB 1802 ATCATTTGATGTGCGAGGCTTGCCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGC 1861
QY 2470 PheGluLeuValGlyAsnThrThrThrThrThrThrThrThrThrThrThrThrThrThr 2489
DB 1862 TTTGAATTTGGGGAAATACTACCCCTTTTGGAGAAATATGTCATCTGCTGCTGAGGA 1921
QY 2490 LysProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPhe 2509
DB 1922 AAACCAACATGTAAGCCATTGAGTGCCTGAAACCCAGAGAGATTTTGAATGGCAATTC 1981
QY 2510 SerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArg 2529
DB 1982 TCTTACACGAGCTTACACTATGACAGCCGTTACTCTTTCGAACCCAGGCGCTTTTCGG 2041
QY 2530 LeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTyrAspValAspAlaPro 2549
DB 2042 CTCGAAGTCCCAAGTCCCTTGACCTGTTTAGAGACAGGTGATTTGGATGTAGATGCCCA 2101
QY 2550 SerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGly 2569
DB 2102 TCTTGCATATGCCATCCACTGTGATTTCCCAACCCATTTGAAATGGTTTTGTAGAGGT 2161
QY 2570 AlaAspTyrSerTyrGlyAlaIleIleTyrSerCysPheProGlyPheGlnValAla 2589
DB 2162 GCAGATTACAGCTATGCTGCCATATCAATCATAGTTGCTTCTCCCTGGGTTTCAGGTGGCT 2221
QY 2590 GlyHisAlaMetGlnThrCysGluGluSerGlyTyrSerSerSerIleProThrCysMet 2609

Db 2222 GGTATGCGCATGAGACCTGTGAAGAGTCAGGATGGTCAAGTTCCATCCCAACATGTATG 2281
Qy |||||
Db 2610 ProIleAspCysGlyLeuProHisIleAapPheGlyAspCysThrIysLeuIysAsp 2629
Qy |||||
Db 2282 CCAATAGACTGTGGCTCCCTCTCATATAGATTTTGGAGACTGTACTAAACTCAAGAT 2341
Qy |||||
Db 2630 AspGlnGlyTrPheGlnGlnGluAspAspMetMetGluValProIleHis 2649
Qy |||||
Db 2342 GACCAAGGATATTTGAGCAAGAGAGACATCATGGAAGTTCCATATGTGACTCTCTCAC 2401
Qy |||||
Db 2650 ProProIleHisLeuGlyAlaValAlaLysThrTrpGluAsnThrIysGluSerProAla 2669
Qy |||||
Db 2402 CCTCCTTATCATTTGGGAGCAGTGGCTAAACCTGGGAAAATACAAAGGAGTCTCCTGCT 2461
Qy |||||
Db 2670 ThrHisSerSerAsnPhenLeuTrpGlyThrMetValSerTrpThrCysAsnProGlyTrp 2689
Qy |||||
Db 2462 ACACATTTCAATCAAACTTTCTGTATGTGTACATGGTTTCTATACACCTGTGTAAATCCAGGATAT 2521
Qy |||||
Db 2690 GluLeuLeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAla 2709
Qy |||||
Db 2522 GAATCTCTGGGAGACCTGTGCTGATCTGCCAGGAGATGGAATCTAGGATGGCAGTGCA 2581
Qy |||||
Db 2710 ProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArg 2729
Qy |||||
Db 2582 CCATCTCTGATTCATTTGAATGTGACTTGCCTACTGCTCTCTGAAAATGGCTTTTGGGT 2641
Qy |||||
Db 2730 PheThrGluThrSerMetGlySerAlaValGlnTrpSerCysIysProGlyHisIleLeu 2749
Qy |||||
Db 2642 TTTACAGAGACTAGCATGGGAGTGTGTGCAATATAGCTGTAAACCTGGACACATTTCTA 2701
Qy |||||
Db 2750 AlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArg 2769
Qy |||||
Db 2702 GTGGGCTCTGGCTTAAGGCTTTGTCTAGAGATAGAAAGTGGAGTGGTGGCTCCCAAGC 2761
Qy |||||
Db 2770 CysGluAlaIleSerCysIysIysProAsnProValMetAsnGlySerIleIysGlySer 2789
Qy |||||
Db 2762 TGTGAAGCCATTTTCATGAAAAAGCCAAATCCAGTTCATGAATGGATCCATCAAGGAAGC 2821
Qy |||||
Db 2790 AsnTrpThrTrpLeuSerThrLeuTrpTrpGluCysAspProGlyTrpValLeuAsnGly 2809
Qy |||||
Db 2822 AACTACACATACCTGAGCAGCTGTGTACTATAGTGTGACCCCGGATATGTGTGTAATGGC 2881
Qy |||||
Db 2810 ThrGluArgArgThrCysGlnAspAspIysAsnTrpAspGluAspProIleCysIle 2829
Qy |||||
Db 2882 ACTGAGAGAGAACATGCCAGATGACAAAACTGGGATGAGATGAGGCCATTTGCATT 2941
Qy |||||
Db 2830 ProValAspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTrp 2849
Qy |||||
Db 2942 CTGTGGACCGCAGTTTCAACCCCGAGTCTCAGCCCAATGGCCAGGTGAGAGACGAGTAC 3001
Qy |||||
Db 2850 ThrPheGlnIysGluIleGluTrpCysAsnGluGlyPheLeuLeuGluGlyAlaArg 2869
Qy |||||
Db 3002 ACATTCCAAAAAGAGATTCGAATACATTCGAATGAAAGGGTTCTTGTGAGGAGCCAGG 3061
Qy |||||
Db 2870 SerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValProVal 2889
Qy |||||
Db 3062 AGTCGGGTTTGTCTTGCCAAATGGAAGTTGGAGTTGGAGTGAGCCACTCCCGACTGTGTGCTGTC 3121
Qy |||||
Db 2890 ArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTrpGlyPhe 2909
Qy |||||
Db 3122 AGATGTGCCACCCCGCCACACTGGCCAAATGGGGTTCAGCGAGGCTTGGAGCTATGGCTTC 3181
Qy |||||
Db 2910 MetIysGluValThrPheHisCysHisGluGlyTrpIleLeuHisGlyAlaProIysLeu 2929
Qy |||||
Db 3182 ATGAGGAAGTAACTTCACATGTACAGAGGGCTACATCTTGCACGGTGTCTCAAAACTC 3241
Qy |||||
Db 2930 ThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysIysProValAsnCys 2949
Qy |||||
Db 3242 ACCTGTCACTCAGATGGCACTGGGATGCAGAGATTTCTCTCTGTAAACCACTCACTCT 3301
Qy |||||
Db 2950 GlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGly 2969
Qy |||||

Db 3302 GGACCTCTCTGAAGACTCTTGCCCATGTGTTCCCTAAATGGTTTCTTTATTTATTCATGGGGC 3361
Qy |||||
Db 2970 HisIleGlnTrpGlnCysPheProGlyTrpLysLeuHisGlyAsnSerSerArgArgCys 2989
Qy |||||
Db 3362 CATATACAGTATCAGTGTCTTCTGTGTATAGCTTCCATGGAAATTCATCAAGAAAGTGC 3421
Qy |||||
Db 2990 LeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThr 3009
Qy |||||
Db 3422 CTCTCCAAATGGCTCTCTGGAGTGGCAGCTACCTTCTCTGCCCTTCCAGATGTTCCACA 3481
Qy |||||
Db 3010 ProValIleGluTrpGlyThrValAsnGlyThrAspPheAspCysGlyIysAlaAlaArg 3029
Qy |||||
Db 3482 CCAGTAAATGAATATGAACTGTCAATGGGACAGATTTTCACTGTGTGAAAGGCGCCGG 3541
Qy |||||
Db 3030 IleGlnCysPheIysGlyPheIysLeuLeuGlyLeuSerGluIleThrCysGluAlaAsp 3049
Qy |||||
Db 3542 ATTCAGTGTCTTAAGGCTTCAAGCTCTTAGGACTTCTGAAATCACCTGTGAAGCCGAT 3601
Qy |||||
Db 3050 GlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMet 3069
Qy |||||
Db 3602 GGCCAGTGGAGCTCTGGTTTCCCCCACTGTGAACACACTTCTTGTGTCTCTTTCCAATG 3661
Qy |||||
Db 3070 IleProAsnAlaPheIleSerGluThrSerSerTrpIysGluAsnValIleThrTrpSer 3089
Qy |||||
Db 3662 ATACCAATCGTTCATCAGTGAGACCCAGCTCTTGGAAAGGAAATGTGATAACTTACAGC 3721
Qy |||||
Db 3090 CysArgSerGlyTrpValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyVal 3109
Qy |||||
Db 3722 TGCAGTCTGGATATGTATACAGGCGATTCAGATCTGATTTGTACAGAGAAAGGGGTA 3781
Qy |||||
Db 3110 TrpSerGlnProTrpValCysGluProLeuSerCysGlySerProSerValAla 3129
Qy |||||
Db 3782 TGGAGCCAGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGTCCCCACCGCTCTGTGCC 3841
Qy |||||
Db 3130 AsnAlaValAlaThrGlyGluAlaHisThrTrpGluSerGluValIysLeuArgCysLeu 3149
Qy |||||
Db 3842 AATGCACTGGCACTGGAGAGGCACACACTATGAAAGTAGAGTGAACCTCAGATGTCTG 3901
Qy |||||
Db 3150 GluGlyTrpThrMetAspThrAspThrAspThrPheThrCysGlnIysAspGlyArgTrp 3169
Qy |||||
Db 3902 GAAGTTATACGATGATACGATACGATACGATACATTCACCTGTTCAGAAAGATGGTCCGTGG 3961
Qy |||||
Db 3170 PheProGluArgIleSerCysSerProLysGlyCysProLeuProGluAsnIleThrHis 3189
Qy |||||
Db 3962 TTCCCTCGAGAGATCTCTGCACTCTTAAATAATGTCTCTCCCGGAAAAACATAACACAT 4021
Qy |||||
Db 3190 IleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGlu 3209
Qy |||||
Db 4022 ATACTTTGTTTCATGGGACGATTTTCAGTGTGAATAGGCAGTTCCTGTGTCTCATGTGCGAA 4081
Qy |||||
Db 3210 GlyTrpThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyTrpTrpGluPro 3229
Qy |||||
Db 4082 GGGTATACCTTTGAGGGAGTTAACTATCAGTATGTGAGCTTGTATGATGACCTGGAGCCA 4141
Qy |||||
Db 3230 ProPheSerAspGluSerCysSerProValSerCysGlyIysProGluSerProGluHis 3249
Qy |||||
Db 4142 CCATTCCTCGATGAATCTTTCAGTCCAGTTTCTTGTGGGAAACCTGAAAGTCCAGAACAT 4201
Qy |||||
Db 3250 GlyPheValIleGlySerIysTrpThrPheGluSerThrIleIleTrpGlnCysGluPro 3269
Qy |||||
Db 4202 GGAATTTGTGGTGGCAGTAAATACCTTTTGAAGCACAGATTTATTTATCAGTGTGAGCCT 4261
Qy |||||
Db 3270 GlyTrpGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGly 3289
Qy |||||
Db 4262 GGCATGAATAGAGGGGAAACAGGAAACGTGTCTGCCAGAGAAACACAGAGTGGAGTGA 4321
Qy |||||
Db 3290 GlyValAlaIleCysIysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyIys 3309
Qy |||||
Db 4322 GGGGTGGCAATATGCAAGAGACCCAGTGTGAAACTCAACTTGAATTTCTCAATGGGAAA 4381
Qy |||||
Db 3310 AlaAsnIleGluAsnArgThrThrGlyProAsnValIleTrpSerCysAsnArgGlyTrp 3329
Qy |||||
Db 4382 GCTGACATTTGAAAAACAGGACGACTGGACCCCAACGTGTGTATATTTCTTCCAAACAGAGGCTAC 4441

QY 3330 SerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProVal 3349
 DB |||||
 DB 4442 AGTCCTGAAGGCCCATCTGGAGCACACTGACAGAAAATGGAACTCTGGAGCCACCCAGTC 4501
 QY 3350 ProLeuGlyLysProAsnProCysProValProPheValIleProGluAsnAlaLeu 3369
 DB |||||
 DB 4502 CCTCTCTGCAACCAATCCATGCCCTGTCTCTTTGTGATGATCCGAGAAATGCTCTGCTG 4561
 QY 3370 SerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeu 3389
 DB |||||
 DB 4562 TCTGAAGAGGAGTTTATGTTGATCAGATGTGTCCATCAATGTAGGGAAGTTTCTG 4621
 QY 3390 LeuGlnGlyHisGlyIleLeuThrCysAsnProAspGluThrTrpThrGlnThrSerAla 3409
 DB |||||
 DB 4622 CTGACGGGCCACCGCATCATTAACCTACACCCGACGACGCTGGACACAGCAAGCGCC 4681
 QY 3410 LysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGly 3429
 DB |||||
 DB 4682 AAATGTGAATAAATCTCATGTGTGTCACCGCTCAGTAGAATAATGCAATTCCTCGAGGC 4741
 QY 3430 ValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGlu 3449
 DB |||||
 DB 4742 GTACATATCAATATGGACATGATCACCTACTCATGTTACAGTGAATGATGTTGGAG 4801
 QY 3450 GlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProPheCysArg 3469
 DB |||||
 DB 4802 GGTTCCTCGAGGAGTGTGTTAGAAAATGGAACATGACATCACCTCTTATTTGCAGA 4861
 QY 3470 AlaValCysArgPheProCysGlnAsnGlyIleCysGlnArgProAsnAlaCysSer 3489
 DB |||||
 DB 4862 GCTGTCTGTGCGATTTCCATGTGAGAAATGGGGCATCTGCAACGCCCAATGCTTGTTC 4921
 QY 3490 CysProGluGlyTrpMetGlyArgLeuGluGluProIleCysIleLeuProCysLeu 3509
 DB |||||
 DB 4922 TGTCCAGAGGCTGATGAGGGGCGCTCTGTGAGAACCAATCTGCATCTTCCCTGCTG 4981
 QY 3510 AsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProGlyTrpThrGlySer 3529
 DB |||||
 DB 4982 AACCGAGGTGCTGTGTGCGCCCTTACAGTGTGACTGCGCCCTGGTGGAGCGGCTCT 5041
 QY 3530 ArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArgPro 3549
 DB |||||
 DB 5042 CGCTGTCTACAGCTGTGTGCGAGCTCTCCCTGCTTAAATGGTGGAAATGTGTAGACCA 5101
 QY 3550 AsnArgCysHisCysLeuSerTrpThrGlyHisAsnCysSerArgLysArgThr 3569
 DB |||||
 DB 5102 AACCGATGTCACTGTCTTCTTCTTGGACGGGACATACCTGTTCCAGGAAGGAGGACT 5161
 QY 3570 GlyPhe 3571
 DB |||||
 DB 5162 GGGTTT 5167

RESULT 7
 ABA03880
 ID ABA03880 standard; cDNA; 6153 BP.

AC ABA03880;
 DT 14-FEB-2002 (first entry)
 XX Human POLY11 nucleotide sequence SEQ ID NO:21.
 XX Human; POLYX; gamma aminobutyric acid receptor; GABA receptor;
 KW epidermal growth factor; EGF; complement receptor; HSPC; syntaxin;
 KW haematopoietic stem and progenitor cell; sulphotransferase; prohibitin;
 KW antidepressant; neuroleptic; antiparkinsonian; norepinephrine; relaxant;
 KW anticonvulsant; neuroleptic; neuroprotective; antialcoholic; cardiant;
 KW tranquilliser; antiarrhythmic; psychiatric; medical; depression; anxiety;
 KW Parkinson's disease; Huntington's disease; Tourette's syndrome; stroke;
 KW amyotrophic lateral sclerosis; head trauma; Alzheimer's disease;
 KW alcoholism; vigilance; muscle tension; epileptogenic; memory function;
 KW cardiomyopathy; arrhythmogenic right ventricular dysplasia; 88.

XX OS Homo sapiens.
 XX Key
 FH CDS
 FT 1. 6153
 FT /tag= a "POLY11"
 FT /product= "POLY11"
 FT /transl_except= (pos:1099..1101,aa:Xaa)
 FT /note= "Xaa is unknown"
 XX WO200179294-A2.
 XX 25-OCT-2001.
 XX 19-APR-2001; 2001WO-US012854.
 XX 19-APR-2000; 2000US-0198293P.
 XX 20-APR-2000; 2000US-0198645P.
 XX 25-APR-2000; 2000US-0199476P.
 XX 26-APR-2000; 2000US-0199880P.
 XX 26-APR-2000; 2000US-0200024P.
 XX 26-APR-2000; 2000US-020025P.
 XX 09-JUN-2000; 2000US-0210809P.
 XX 17-JUL-2000; 2000US-0218591P.
 XX 11-AUG-2000; 2000US-0224610P.
 XX 09-FEB-2001; 2000US-0267673P.
 XX 27-FEB-2001; 2001US-0271814P.
 XX (CURA-) CURAGEN CORP.
 XX Taupier RJ, Vernet CAM, Fernandes E, Shinkets RA, Majumder K;
 PI Padigaru M, Colman SD, Zerhuseen BD, Spytek KA, Burgess CE, Liu X;
 XX WPI; 2002-017601/02.
 XX P-PSDB; AAG68264.
 XX New isolated polypeptides for treating a broad range of pathological
 PT states, e.g., depression, stroke, Parkinson's disease, Huntington's
 PT disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
 PT and Alzheimer's.
 XX Claim 9; Page 43-45; 155pp; English.
 XX The present invention describes polypeptides (I), designated POLYX
 CC polypeptides (e.g. POLY1, POLY2, etc.), and the polynucleotide sequences
 CC (II) encoding them. POLY1-4 are members of the gamma aminobutyric acid
 CC (GABA) receptor family; POLY5-8 are members of the epidermal growth
 CC factor (EGF) family; POLY9-11 are members of the complement receptor
 CC family; POLY12 is a member of the haematopoietic stem and progenitor cell
 CC (HSPC) family; POLY13 is a member of the sulphotransferase family; POLY14
 CC -16 are members of the syntaxin family; and POLY17 is a member of the
 CC prohibitin family. (I) and (II) can have antidepressant,
 CC neuroprotective, antiparkinsonian, norepinephrine, relaxant, anticonvulsant,
 CC neuroleptic, neuroprotective, antialcoholic, cardiant, tranquiliser and
 CC antiarrhythmic activities. (I) and (II) can be used for treating or
 CC preventing a POLYX-associated disorder in humans as a therapeutic in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from a POLYX-associated disorder, for treating a
 CC pathological state in a mammal, especially patients suffering from, e.g.,
 CC psychiatric and medical conditions, depression, stroke, Parkinson's
 CC disease, Huntington's disease, Tourette's syndrome, amyotrophic lateral
 CC sclerosis, head trauma, Alzheimer's disease, alcoholism, vigilance,
 CC anxiety, muscle tension, epileptogenic activity and memory functions,
 CC cardiomyopathy and arrhythmogenic right ventricular dysplasia. The
 CC present sequence encodes POLY11
 XX Sequence 6153 BP; 1610 A; 1482 C; 1563 G; 1497 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 0 Length: 6153
 Score: 8427.00 Matches: 1535
 Percent Similarity: 78.65% Conservative: 1
 Best Local Similarity: 78.60% Mismatches: 5

Query Match: 42.19% Indels: 412
DB: 6 Gaps: 3
US-09-977-053-4 (1-3571) x ABA03880 (1-6153)

1993 TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSer 2012
193 TATACTCTTGCTGGTCTTGACACCATTTGAATGCTGGCCGACGGCAAGTGGAGTAGAGT 252
2013 AspGlnGlnCysLeuAlaValSerCysAspGluProProProlleValAlaAspHisAlaSerPro 2032
253 GACCAGCAGTGCCTGCTGCTCTGATGAGCCACCCATTTGTCGACCAACGAGCTCTCCA 312
2033 GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSer 2052
313 GAGACTGCCCATGGCTCTTTGGAGACATTTGCATTTACTACTGCTCTGATGTTTACAGC 372
2053 LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGly 2072
373 CTAGCAGACATTTCCAGCTTCTCTGCAATGCCCAGGSCAAGTGGGTACCCCCAGAGGT 432
2073 GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer 2092
433 CAAGACATGCCCGTGTGTATAGCTCATTTCTGTGAAAAACCTCCATCGGTTCCTATAGC 492
2093 IleLeuGlnSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys 2112
493 ATCTTGGAACTGTGAGCAAGCAAAATTTGCAAGCTGGCTCAGTTGTGAGCTTTAAATGC 552
2113 MetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrp 2132
553 ATGGAAGGCTTTGTACTGAACCTCAGCAAGATTTGAATGATGAGAGTGGCGAGTGG 612
2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle 2152
613 AACCCCTTCCCATGATGCCATCCATCCCTGTGCGGTGTGGAGAGCCACCAAGCATC 672
2153 MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
673 ATGAATGGCTATGCAAGTGGATCAACTACAGTTTGGAGCCATGTTGGCTTACAGCTGC 732
2173 AsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThrGlyGlnTrp 2192
733 AACAGGGGTTCTACATCAAGGGGAAAAAGAGAGCACCTGCGAAGCCACAGGGCAGTGG 792
2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
793 AGTAGTCTTATACCGACGTGCCACCCGGTATCTTGTGGTGAACACCATTAAGGTTGAGAAT 852
2213 GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn 2232
853 GGCTTTCTGGAGATACAACTGTCAGAGTCTTTGAGAGTGAAGTGAAGTATCAGTGTAAAC 912
2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHis 2252
913 CCGGGCTATAAGTCAGTCGGAGTCTGTATTTGTCTGCCAAGCCATCGCCACTGGCAC 972
2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProProlleGlnAsn 2272
973 AGTGAATCCCTCTGATGTGTGTCTCTCGACTGTGGAAAAACCTCCCGGATCCAGAAAT 1032
2273 GlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
1033 GGCTTCATGAAGAGGAGAAAATTTGAAAGTAGGGTCCAGGGTCAAGTTTCTGTAAATGAA 1092
2293 Gly----TyrGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsn 2311
1093 GGTTTATNGAGCTTTGTTGGGACAGTCTTGGACATGTGAGAAATCTGCAAAATGGAAT 1152
2312 LysLysSerAsnProLysCysMetProAlaLysCysPheProGluProProLeuLeuGluAsn 2331
1153 AAGAAGTCAAAATCCAAAGTGCATGCTGCAAGTGGCCAGAGCCGCCCTCTTGGAAAC 1212

2332 GlnLeuValLeuLysGluLeuThrThrGluValValGlyValValThrPheSerCysLysGlu 2351
1213 CAGCTAGTATTAAAGGAGTTGACCAACGAGGTAGGAGTTGTGACATTTCTCTTAAGAA 1272
2352 GlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAsp 2371
1273 AGGCATGTCCTCGCAAGGCCCTCTGCTCTGAAATGCTTGGCATCCAGCAATGGAATGAC 1332
2372 SerPheProValCysLysIleValLeuLysThrProProProlleIleSerPheGlyVal 2391
1333 TCTTTCCCTCTTTGTAAGATTGTTCTTTGTACCCCACTCCCTTAATTTCTCTTGGGTGC 1392
2392 ProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGly 2411
1393 CCATTTCTTCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1452
2412 PhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerPro 2431
1453 TTTTCTCTTAAGAGAAATTTCTACACCTCTGCAACCTGATGGCAGCTGGAGCTCTCCA 1512
2432 LeuProGluCysValProValGluCysProGluProGluGluIleProAsnGlyIleIle 2451
1513 CTGCCAGAAATGTTCCAGTAGAATGTTCCCAACCTGAGGAAATCCCAATGGAATCAAT 1572
2452 AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGlu 2471
1573 GATGTGCAAGGCTTGGCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGCTTTGAA 1632
2472 LeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysPro 2491
1633 TTGGTGGGAAATACTACCACTTTGTGGAGAAAAATGGTCACTGGCTTGGAGGAAACCA 1692
2492 ThrCysLysAlaIleGluCysLysLysProLysGluIleLeuAsnGlyLysPheSerTyr 2511
1693 ACATGTAAAGCCATTCAGTGCCTGAAACCCCAAGAGATTTTGAATGGCAATTTCTTTAC 1752
2512 ThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGlu 2531
1753 ACGGACTACACTATGGACGACGCTTACTCTCTTGTCAACCGAGGCTTTCGGCTCGAA 1812
2532 GlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCys 2551
1813 GGTCCCAAGTGCCTTGACCTGTTTAGAGACAGGTGATTTGGGATGTAGATGCCCATCTTC 1872
2552 AsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAsp 2571
1873 AATGCCATCCATGTGATTTCCCAACCCATTTGAATAATGGTTTGTAGAGGTCAGAT 1932
2572 TyrSerTyrGlyAlaIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHis 2591
1933 TACAGCTATGGTGGCATATCATCTACAGTTGCTTCCCTGGGTTTCAGGTGGCTGGTCAT 1992
2592 AlaMetGlnThrCysGluGluSerGlyTrpSerSerSerIleProThrCysMetProlle 2611
1993 GCCATGCAGACCTGTGAAGAGTCAAGTGTCAAGTTCCATCCCAATGTATGCCAATA 2052
2612 AspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGln 2631
2053 GACTGTGGGCTCCCTCTCATATAGATTTTGGAGACTGTACTAACTCAAGATGACCAAG 2112
2632 GlyTyrPheGluGlnGluAspMetMetGluValProTyrValThrProHisProPro 2651
2113 GGATATTTTGGACAGACAGACATGATGGAAGTTCCATATGTGACTCTCTCACCCCTCT 2172
2652 TyrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHis 2671
2173 TATCATTTGGAGCAGTGGCTAAACCTGGGAAAAATACAAAGAGGTCTCTCTGTCTACAT 2232
2672 SerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeu 2691
2233 TCATCAAACTTTCTGTGTGTACCATGGTTTCATACACCTGTATATCCAGGATATGAACCT 2292
2692 LeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSer 2711

2293 CTTGGGAAACCTCTGTGTGATCTGCCAGGAAGATGGAACTTGGAAATGGCAGTGCACCATCC 2352
2712 CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
2353 TGCATTTCAATTTGAATGTGACTTGTCTACTCTCTGCTGAAATGGCTTTTGGTTTTACA 2412
2732 GluThrSerMetGlySerAlaValGlnTyrSerCysIleProGlyHisIleLeuAlaGly 2751
2413 GAGACTAGCATGGGAAGTCTGTGCAATATAGCTGTAACCTGGACACATTTCTAGCAGGC 2472
2752 SerAspLeuArgLeuCysLeuGluAsnArgIleTyrSerGluValAspProArgCysGlu 2771
2473 TCTGACTTAAGGCTTTGTCTAGAGAATAGAAAGTGGAGTGGTCCCTCCCAAGCTGTGAA 2532
2772 AlaIleSerCysIleSerProAsnProValMetAsnGlySerIleGlySerAsnTyr 2791
2533 GCCATTTTCATCAAAAAGCCAAATCCAGTCAGTGAATGGATCCATCAAGAGCAACTAC 2592
2792 ThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGlu 2811
2593 ACATACCTGAGCAGCTGTACTATGAGTGTGACCCCGATATGTGCTGAATGGCACTGAG 2652
2812 ArgArgThrCysGlnAspAspIleValAsnGluAspGluProIleCysIleProVal 2831
2653 AGGAGAACATGCCAGGATGACAAAACTGGGATGAGGATGAGCCCAATTTGCATTTCTGTG 2712
2832 AspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPhe 2851
2713 GACTGCAGTTCAACCCCTCAGCCCAATGCCAGTGCAGAGGAGAGACAGTACACATTC 2772
2852 GlnIleGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArg 2871
2773 CAAAAGAGATGGAATACACTTGCATGAAGGGTCTTCTGCTGAGGGAGCCAGAGTCGG 2832
2872 ValCysLeuAlaAsnGlySerTyrSerGlyAlaThrProAspCysValProValArgCys 2891
2833 GTTGTCTTGGCAATGAAGTGGAGTGGAGCCACTCCGACTGTGTGCTGTGCATGATGT 2892
2892 AlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLys 2911
2893 GCCACCCGCCCAACATGGCCCAATGGGTGACGGAAGGCTGGACTATGGCTTCATGAAG 2952
2912 GluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProIleLeuThrCys 2931
2953 GAAGTAACATTCACCTGTACGAGGGCTACATCTTGCACGGTGTCTCCAAAACCTCACCTGT 3012
2932 GlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysIleSerProValAsnCysGlyPro 2951
3013 CAGTCAGATGGCACTGGGATGAGAGATTCCTCTCTGTAACACCACTCAACTGTGGACCT 3072
2952 ProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIle 2971
3073 CCTGAAGATCTTGCCTTCCCTAATGGTTTTCTTCTTATTATTCATGGGGCCATATA 3132
2972 GlnTyrGlnCysPheProGlyTyrIleLeuHisGlyAsnSerSerArgArgCysLeuSer 2991
3133 CAGTATCAGTGTCTTCTGCTTATAAGCTCCATGGAAATTCATCAAGAAGTGCCTCTCC 3192
2992 AsnGlySerTrpSerGlySerProSerCysLeuProCysArgCysSerThrProVal 3011
3193 AATGGCTCTGGAGTGGCAGCTCACCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCT 3252
3012 IleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyIleValAlaArgIleGln 3031
3253 ATTGAATATGGAATCTGCAATGGACAGATTTTGTGCTGGAAAGCAGCCCGGATTCAG 3312
3032 CysPheIleGlyPheIleLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln 3051
3313 TGCTTCAAAAGGCTTCAAGCTCTTAGGACTTTCTGAAATATCACTGTGAAGCCGATGCCAG 3372
3052 TrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIlePro 3071

3373 TGGAGCTCTGGGTTCCCCCACTGTGAACACACTTCTTCTGTGGTTCTCTTCCAAATGATACCA 3432
3072 AsnAlaPheIleSerGluThrSerSerTrpLysGluAsnValIleThrTyrSerCysArg 3091
3433 AATGCGTTTCATCAGTGCAGACCACTCTTGGAGGAAAAATGTGATAACTTACAGCTGCAGG 3492
3092 SerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSer 3111
3493 TCTGGATATGTATACAGGCAGTTCAGATCTGATTTGTACAGAGAAAGGGGTATGGAGC 3552
3112 GlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAlaAsnAla 3131
3553 CAGCCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCCCAACCGTCTGTGCGCAATGCA 3612
3132 ValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGly 3151
3613 GTGGCAACTGGAGAGGACACACCTATGAAAGTGAAGTGAAGTCAAGATCTCAGATGCTCTGGAAGGT 3672
3152 TyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPhePro 3171
3673 TATACGATGGATACAGATACAGATACATTCACCTGTGCAGAAAGATGGTCTGCTGCTCCCT 3732
3172 GluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeu 3191
3733 GAGAGATCTCTGCGAGTCTTAAAAATGTCTCTCCGGAANAACATACACATATACTT 3752
3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
3793 GTACATGGGAGCGATTTTCAGTGTGAATGAGCAAGTTTCTGTGTCATGTGCAGAGGGTAT 3852
3212 ThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProPhe 3231
3853 ACCTTTGAGGGAGTTAAACATATCAGTATGTACGCTTGATGAACCTGGAGGCCACCATTC 3912
3232 SerAspGluSerCysSerProValSerCysGlyLysProGluSerProGluHisGlyPhe 3251
3913 TCCGATGATCTTGCAGTCCAGTTTCTGTGGGAACCTGGAAGTCCAGAACATGGATTT 3972
3252 ValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyr 3271
3973 GTGGTTGGCAGTAAATACACCTTTGAAAGCACAAATTTATTTATCAGTGTGAGCCCTGGCTAT 4032
3272 GluLeuGlu----- 3274
4033 GAACATAGAAATTTGGCTGTGAATCCATCTGGTCTGGACTTTCTTGTGTGACAGGACC 4092
3274 ----- 3274
4093 CTCAGCTGCAGGTTCGGAGTTGGCTAGAGGTCCAATCCAGACCCTGTTTGTGCTGGGTATCA 4152
3274 ----- 3274
4153 GCAGCAGAGGGTGCAGAACAGCGGATATTGGTGAACCGCAAAATGCTGCTGCTGCTGATCAT 4212
3274 ----- 3274
4213 CTTCTGGAAGTTTGTCTCAGAGGAATACCCGGCCATGTGAGGTGTGAGTCCGCCCTAC 4272
3275 -----GlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyIleValAla 3292
4273 TGGGGGGGGAACAGGGAACGCTGTCCAGAGAGAACAGACAGTGGAGTGGAGGGTGGCA 4332
3293 IleCysGlyGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIle 3312
4333 ATATGCAAGAGACCAAGGTGTGAACCTCCACTTGAATTTCTCAATGGGAAGCTGACAT 4392
3313 GluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGlu 3332
4393 GAAAAACAGACGACTGGAACCAACGTGTATATTTCTGCAACAGAGGCTACAGCTTGTAA 4452
3333 GlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCys 3352
4453 GGCCCATCTGAGGCACACTGCACAGAAAAATGGAACCTGGAGGCCACCCAGTCCCTCTCTGC 4512

```

QY 3353 LysProAsnProCysProVal1ProPheVal1LeProGluAsnAlaLeuLeuSerGluLys 3372
DB 4513 AAACCAATTCATGCCCTGTTCTTTGTATGATCCCGAAGATGCTGCTGTGAAAG 4572
QY 3373 GluPheTyrVal1AspGluAsnVal1Ser1LeuLysCysArgGluGlyPheLeuLeuGlnGly 3392
DB 4573 GAGTTTATGTTATCATGAATGTCATCAATGATGAGGAGGTTTCTGCTCAGGAC 4632
QY 3393 HisGluTyrLe1LeuThrCysAsnProAspGluThrTrpThrGluThrSerAlaLysCysGlu 3412
DB 4633 CACGGATCATATCTGCAACCCGACGACGTGACACAGACAGGCCAAATGTGA 4692
QY 3413 Lys----- 3413
DB 4693 AGAAGATATACACACAGCCCAAGTCCCTGATTTTTCAGTACGAGCTTATTGAGTAT 4752
QY 3413 ----- 3413
DB 4753 AGAATGTTATTTTGCAGGAGGAGGAGTTCAAGATGACCAATGAAACAGCTGTGACGGA 4812
QY 3413 ----- 3413
DB 4813 GCCTCCACCGAGAGAACAAAAACAAAGCGAGAAAAGCAGGTGTACAGGCCCA 4872
QY 3413 ----- 3413
DB 4873 CTTGGAGCAACATGGGGCAAGCAGAGCTCCACCCCAAGCAAGAGGTGACCTCC 4932
QY 3413 ----- 3413
DB 4933 TGGGGAAATTCAGCAACTCCAGCGAGGATTATGACAGACCTGTGATCTCCTGAGA 4992
QY 3413 ----- 3413
DB 4993 TGGAGCCCTGGGCTCATATGTCATGTCCTCAGACAGATCAGAGCTTATGCTTCC 5052
QY 3413 ----- 3413
DB 5053 CTTGCTGCTCTGAGAAATCCAGGAGGCTGAGCTAGTGGGATTCCTCCACAGCAGTTT 5112
QY 3413 ----- 3413
DB 5113 ACCTGCTCTGCAAGGGGCAAGCTAGAGCCTTTGTAAGCGATCCCTGATCCATGCT 5172
QY 3413 ----- 3413
DB 5173 CTTGATTTGGATAGACCCCCCAACAGGGGTCAAGGATGAAACCCCCCAACAGG 5232
QY 3413 ----- 3413
DB 5233 GGTACACAGACACTTATACAGGGGTGTTCTGTGATGATCAGGTGAGTCCCTGAGG 5292
QY 3413 ----- 3413
DB 5293 ACAGAGCTCCAGAGAAAGAGCAGGACCATCTTGTGCTGTGACGCTCCGTGGA 5352
QY 3413 ----- 3413
DB 5353 AAAGCAGCAATTTGGGAGAGGCTTAGATTGATGAATTTGAAAGATGATGCTTCAGAAAG 5412
QY 3413 ----- 3413
DB 5413 TGGGTATATATGAATTCGTAGCTTAAGAGACATGTTCTTAACCAATGCAAGAGCC 5472
QY 3413 ----- 3413
DB 5473 AAGAACAGAGATTAACATTACAGATCCGTTAACAGAAATMACAGTTTAGAAAGAT 5532
QY 3413 ----- 3413
DB 5533 GTAAATGACCTGATGAGCTGAAAAACACACAGAACTTTCATATGACACACAAAA 5592

```

```

QY 3413 ----- 3413
DB 5593 CAAGGCCAATTCACATTCCAGAAATCCAGAGAACCCAGTAAATATCTCCATGAGAG 5652
QY 3414 ----- 3414
DB 5653 ATCAACCCCAAGACATATATCTCAGGTTCTCCAGAAATCTCATGTGTCACAGCT 5712
QY 3421 HisValGluAsnAla1LeuArgGlyValHisTyrGlnTyrGlyAspMet1LeuThrTyr 3440
DB 5713 CAGTGAAGAAATGCAATTCGTGAGGCTGATCATATATGAGACATGATCACTTAC 5772
QY 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluGlnGly 3460
DB 5773 TCATGTTACAGTGATATCATGTGAGAGGTTCTCCAGAGAGTGTGTTTAAATAATGGA 5832
QY 3461 ThrTrpThrSerProPro1LeuCyArgAlaValCysArgPheProCysGluAsnGlyGly 3480
DB 5833 ACATGACATCACTCTTATTTGACAGACTGTCTGTCAATTCATATGTCAGAAATGAGGAC 5892
QY 3481 IleCysGluArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGlu 3500
DB 5893 ATTCGCCAAGCCCAATGCTTGTTCCTGTCAGAGGCTGATGGGGGCTCTGTGTA 5952
QY 3501 GluProIleCysValLeuProCysLeuGlnGlyValArgCysValAlaProTyrGlnCys 3520
DB 5953 GAACCAATCTGCATCTTCCCTGTCGAAACGAGGCTCCTGTGTGGCCCTTACAGATGT 6012
QY 3521 AspCysProProGlyTyrTrpThrGlySerArgCysHisThr 3533
DB 6013 GACTGCCCTGCTGAGAGGAGGCTGCTGTGTATACA 6051

RESULT 8
ID ABX56476
ABX56476 standard; cDNA; 6152 BP.
AC ABX56476;
XX
DT 17-FEB-2003 (first entry)
DE
XX cDNA encoding novel human complement receptor-like protein #3.
XX
XX Gamma-aminobutyric acid receptor-like protein; depression; stroke;
XX GABA receptor-like protein; Parkinson's disease; Huntington's disease;
XX Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;
XX Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;
XX epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;
XX arrhythmogenic right ventricular dysplasia; renal disease; diabetes;
XX epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;
XX haematopoietic stem and progenitor cell like protein; cirrhosis;
XX sulfoxtransferase-like protein; cholangitis; hepatitis; hyperthyroidism;
XX developmental disorder; Syntaxin-like protein; myxoid liposarcoma;
XX asthma; Lambert-Radon myasthenic syndrome; acute myeloid leukaemia;
XX transgenic animal; gene; SS.
XX
XX Homo sapiens.
XX
XX US2002123612-A1.
XX
XX 05-SEP-2002.
XX
XX 03-JUL-2001; 2001US-00898570.
XX
XX 19-APR-2000; 2000US-0198293P.
XX 20-APR-2000; 2000US-0198645P.
XX 25-APR-2000; 2000US-0199476P.
XX 26-APR-2000; 2000US-0199898P.
XX 26-APR-2000; 2000US-0200024P.
XX 26-APR-2000; 2000US-0200025P.
XX 09-JUN-2000; 2000US-0210809P.
XX 03-JUL-2000; 2000US-0215855P.
XX 17-JUL-2000; 2000US-0218591P.
XX 11-AUG-2000; 2000US-0224610P.

```

PR 27-FEB-2001; 2001US-0271814P.
XX (GERL/) GERLACH V L.
PA (EILE/) EILBERMAN K. L.
PA (MACD/) MACDOUGALL J R.
PA (SMIT/) SMITHSON G.
XX Gerlach VL, Eilberman K, Macdougall JR, Smithson G,
XX MPI: 2003-066815/06.
DR P-PSDB; ABDU2097.
XX
XX Novel polypeptides and nucleic acids which are members of epidermal
PT growth factor, complement receptor families for diagnosing and treating
PT psychiatric conditions, depression, stroke, Alzheimer's and Parkinson's
PT disease.
XX
XX Claim 9; Page 28-31; 91pp; English.
XX
XX The invention describes an isolated POLYX (POLYX-17) polypeptide and its
CC variant. POLYX polypeptides (especially POLY5, POLY6 and POLY7), the
CC polynucleotides encoding them (I) and an anti-POLYX-antibody (III) are
CC useful for treating or preventing a pathology associated with POLYX
CC polypeptide in humans and for treating a syndrome associated with human
CC disease. POLYX polypeptide is also useful for identifying an agent that
CC binds to POLYX and a cell expressing POLYX is useful for identifying a
CC therapeutic agent for use in treatment of a pathology related to aberrant
CC expression or physiological interactions of the polypeptide. (III) is
CC useful for treating a pathological state in a mammal and for determining
CC the presence or amount of POLYX in a sample. POLYX-4 (GABA receptor-like
CC proteins) are useful for the treatment of psychiatric and medical
CC conditions, depression, stroke, Parkinson's disease, Huntington's
CC disease, Tourette's syndrome, amyotrophic lateral sclerosis, head trauma,
CC Alzheimer's disease, alcoholism, vigilance, anxiety, muscle tension,
CC epileptogenic activity and memory functions, cardiomyopathy and
CC arrhythmogenic right ventricular dysplasia. POLY5-8 (Epidermal growth
CC factor like proteins) may be useful for treating cancer, aberrant
CC angiogenesis, renal disease and diabetes. POLY12 (hematopoietic stem and
CC progenitor cell like protein) may be useful for treatment of leukemia,
CC lupus and anaemia. POLY13 (nucleotransferase-like protein) may be useful
CC for treating cirrhosis, cholangitis, hepatitis, ulcers, hyperthyroidism
CC and developmental disorders. POLY14-16 (Synactin-like proteins) may be
CC useful in treatment of Lambert-Radon myasthenic syndrome, asthma, myxoid
CC liposarcoma and acute myeloid leukemia, and POLY 18 may be useful in
CC treatment of cancers. Cells comprising (I) are useful for producing non-
CC human transgenic animals which are useful for studying the function
CC and/or activity of POLYX protein and for identifying and/or evaluating
CC modulators of POLYX protein activity. This sequence encodes a novel human
CC protein
CC
SQ Sequence 6152 BP; 1609 A; 1483 C; 1562 G; 1497 T; 0 U; 1 Other:
Alignment Scores:
Pred. No.: 0 Length: 6152
Score: 8400.50 Matches: 1534
Percent Similarity: 78.55% Conservative: 0
Best Local Similarity: 78.55% Mismatches: 7
Query Match: 42.06% Indels: 413
DB: 7 Gaps: 3
US-09-977-053-4 (1-3571) X ABX56476 (1-6152)
QY 1993 TTTThleuAlaGlyLeuAspThrIleGluCysleuAlaAspGlyLysTrpSerArgSer 2012
DB 193 TATACCTTGGCTGCTGCTTGAACACATTGATTCCTGCGCACGACAGTGAAGTAAAGT 252
QY 2013 AspGlnGlnCysleuAlaValSerCysAspGluProProIleValAlaPheIleAlaSerPro 2032
DB 253 GACCAACAGGCGCTGGCTGCTCTCTGATGAGCCACCATTTGGACACAGCCTCTCCA 312
QY 2033 GluThrAlaIleHisArgLeuPheGlyAspIleAlaPheIleTyrCysSerArgGlyTyrSer 2052
DB 313 GAGACAGCCCATCGGCTCTTGAAGACATTCATCTACTAGCTGATGGTTACAGC 372

QY 2053 IeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnIleGlyTrpValProProGlnGly 2072
DB 373 CTAGACAGACAAATCCCGAGCTTCTCTGCAATGCCAGGCGCAAGTGGTACCCCGAAGGT 432
QY 2073 GlnAspMetProArgCysIleAlaHisPheCysGlnLysProProSerValSerTyrSer 2092
DB 433 CAAGACATGCCCGGTGTATAGCTCATCTTGTGAAAAAACCCTCATGGTTCTATAGC 492
QY 2093 IleLeuGlnSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys 2112
DB 493 ATCTGGAACTCTGTAGCCAAAGCAAAATTTGCAGCTGGCTCAGTTGAGCTTTAAATGC 552
QY 2113 MetGlnGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlnTrp 2132
DB 553 ATGGAAGGCTTTGTACTGAAACCTCAGCAAGAAATGTAATGACAGGGTGGCAGTGG 612
QY 2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle 2152
DB 613 AACCTTCCCGCATGTCATCCAGTCCAGTGCATCCCTGTGGGTGGAGAGCCACAGCATC 672
QY 2153 MetAsnGlyTyrAlaSerGlnSerAsnTyrSerPheGlnValAlaTyrSerCys 2172
DB 673 ATGAATGGCTATGCAAGTGAATCAAACTAAGTTTGAGCCATGATGGCTTACAGCTGC 732
QY 2173 AsnLysGlyPheTyrIleLysGlyGluLysSerThrCysGlnAlaIleArgIleGlnTrp 2192
DB 733 AACAGGGGTTTCATCANTCAAAAGGGAABAGAGACCTGCCAGGCCACAGGGCAGTGG 792
QY 2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
DB 793 AGTAGTCTCTATACCGACGTCCACCCCGTATCTGTGTGTAACCACTTAAGTTGAGAT 852
QY 2213 GlyPheLeuGlnIleIleThrGlnArgIlePheGlnSerGluValArgTyrGlnCysAsn 2232
DB 853 GGCTTTTGAGACATACACTGGGAGATCTTGAAGTGAAGTACAGTATCACTGATAC 912
QY 2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTyrHis 2252
DB 913 CCGGGCATTAAGTCAGTCGGAAGTCCGTATTTGTCGCCAAGCAATCCGACATGGCAC 972
QY 2253 SerGlnSerProLeuMetCysValProLeuAspCysGlyLysProProProIleGlnAsn 2272
DB 973 AGTGAATCCCTCTGAGTGTGTCTCTGACATGTGAAACCTCCCGATCCAGAT 1032
QY 2273 GlyPheMetLysGlnLysAsnPheGlnValGlySerLysValGlnPhePheCysAsnGlu 2292
DB 1033 GGCTTCATGAAGAGAAACCTTGAAGTGAAGTCCAGGCTCATGTTCTGTAAATGA 1092
QY 2293 GLY---TyrGlnLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyrAsn 2311
DB 1093 GGGTTATNAGAGCTTTGTTGGGAGACAGTCTTGACATGTCAAGAAATCGGCAAAATGGAAT 1152
QY 2312 LysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsn 2331
DB 1153 AAGAAAGCAATCAAGATGCATGTCCGCAAGGCCACAGCGCCCTCTTGTGAAAAC 1212
QY 2332 GlnLeuValLeuLysGlnLysLeuThrThrGlnValGlyValIleThrPheSerCysLysGlu 2351
DB 1213 CAGCTGATATTAAAGAGTTGACCAACGAGTGAAGTTGTGACATTTCTCTGTAAGAA 1272
QY 2352 GlyIleValLeuGlnGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAsp 2371
DB 1273 AGGCAATGCTGCAAGGCCCCCTGCTGCTGAATATGCTGCAATCCAGCAATGAAGAAC 1332
QY 2372 SerPheProValCysValIleValLeuCysThrProProProLeuIleSerPheGlyVal 2391
DB 1333 TCTTCCCTTGTGTGAATGTCTTTGTAAACCACTCCCTCAATTTCTTTGGGTGC 1392
QY 2392 ProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGly 2411
DB 1393 CCAATTCCTTCTCTGCTCTTCAATTTTGAAGTACTGCAAGTATCTTGTGTAGTGGG 1452

QY	2412	PhePheLeuAlaGlyIaSerThrThrLeuCysGlnProAlaGlyIThrTrpSerSerPro	2421
Db	1453	TTTTTCTTAAGAGAAATTCTACCAACCTCTGCCAACCTGATGGCACTGGAGCTTCCA	1512
QY	2432	LeuProGluCysValProValGluCysProGlnProGluGluIleProAlaGlyIleIle	2451
Db	1513	CTGCCAGATATGTCATGATGATGTCCTCAACCTGAGAAATCCCATGATCATTT	1572
QY	2452	AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyIpheGlu	2471
Db	1573	GATGTGCAGAGCCCTTGCTATCTACAGCAAGCTCTATATCTGCAAGCCAGGCTTGAA	1632
QY	2472	LeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyIleTrpPheGlyGlyLysPro	2491
Db	1633	TTGGTGGGAAATACTACCAACCTTTGTGGAGAAATGTCTCACTGGCTTGGAGAAACCA	1692
QY	2492	ThrCysLysValAlaGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyr	2511
Db	1693	ACATGTAAAGCCATTGAGTGGCTGAAACCCAGAGATTTGATGGCAAAATCTCTTAC	1752
QY	2512	ThrAspLeuAlaTyrGlyGlnThrValThrTyrSerCysAsnArgGlyIpheArgLeuGlu	2531
Db	1753	ACGAGCCTACACTATGAGACAGACCGTTACTACTCTTGCAACCGAGCGCTTGGGTGAA	1812
QY	2532	GlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCys	2551
Db	1813	GGTCCCAAGTCCCTTGAACCTGTATTAGACAGGTGATGGAGATGTAGAGCCCACTTGC	1872
QY	2552	AsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyIpheValGluGlyIleAsp	2571
Db	1873	AATGCCATCCACTGTGATTTCCCAACCCATGAAATGTGTTTGTAAGAGTGCAGAT	1932
QY	2572	TyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyIpheGlnValAlaGlyHis	2591
Db	1933	TACAGCTATGGTCCATATACATCTACAGTTCCTTCCGTGGTTCAAGTGGCTGGTAT	1992
QY	2592	AlaMetGlnThrCysGluGluSerGlyTrpSerSerIleProThrCysMetProIle	2611
Db	1993	GCCATGACAGACCTGTGAAGAGTCAAGATGTCTCAAGTCCATCCCAACATGTATGCCAATA	2052
QY	2612	AspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGln	2631
Db	2053	GACTGTGACCTCCCTCCTCATATACATTTTGGAGCTGTACTAACTAAAGATACAG	2112
QY	2632	GlyTyrPheGluGlnGluAspAspMetMetGluValProTyrValThrProHisProPro	2651
Db	2113	GGATATTTTGAACAGAAAGACACATGATGAAATTCATATGTACTCTCCACCCCTCT	2172
QY	2652	TyrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHis	2671
Db	2173	TATCATTTTGGAGGAGTGGCTTAAACTGGGAAATATACAAAGAGATCTCTGTCTACAT	2232
QY	2672	SerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyItyrGluLeu	2691
Db	2233	TCATCAAACTTCTGTATGTATGCATATGTTTCATCAACCTGTAAACAGATATCAACTT	2292
QY	2692	LeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSer	2711
Db	2293	CTGGGGAACCCGTGTCTGATCTGCCAGGAAGATGGAACTGGAAATGGCATGGCAACATCC	2352
QY	2712	CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyIpheLeuArgPheThr	2731
Db	2353	TGCATTTTCAATTGAATGTGACTTGGCTACTGTCTCCGAAAAATGGCTTTTGGCTTTTACA	2412
QY	2732	GluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyIHisIleLeuAlaGly	2751
Db	2413	GAGACTAGCATGGGAAGTGTGTGCAGATATACTGTAAACCTGGACACATTTCTACAGAC	2472
QY	2752	SerAspLeuArgLeuCysLeuGluLubAsnArgLysTrpSerGlyAlaSerProArgCysGlu	2771
Db	2473	TCTGACTTAAGCTTTGTCTTACAGATTAAGATGAGTGGTGTGCTCCCAACGCTGTATA	2533
QY	2772	AlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyr	2791

Db	2533	GCATTTCATGCAAAAAGCCAAATTCAGTATGATGATTCATCAAGAGCACTAC	2592
Qy	2792	ThrTYrLeuSerThrIleuTYrTgUCyAspProGlyTYrValIleAsnGlyThrGlu	2811
Db	2593	ACAACTGACGACGGTGTGACTATGAGTGAACCCGGATATGTGGTAATGGCACTGAG	2652
Qy	2812	ArgArgThrCysGlnAspAspIlyAsnTYrAspGluAspGluProIleCYsIleProVal	2831
Db	2653	AGGAGAACAGGCCAGATGACAAAACCTGGGATGAGATGAGCCATTTCGATTCCTGTG	2712
Qy	2832	AspCysSerSerProProValSerAlaAsnGlyGluValArgGlyAspGluTYrThrPhe	2851
Db	2713	GACTGCACTTACCCCAAGCTCAAGCCAAAGGCGAGGTGAGAGAGACGATACACATTC	2772
Qy	2852	GlnIlyGlnIleGluTYrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArg	2871
Db	2773	CAAAAAGAGTTGATATACATTCGAAAGAGGGTTCCTTGATGAGGAGCCAGAGTCCG	2832
Qy	2872	ValCysLeuAlaAsnGlySerTrpSerGlyValAlaThrProAspCysValProValArgCys	2891
Db	2833	GTTGTGCTTCCAATGAGATGGATGGAGTGGAGCCACTCCCACTGTGTGCTTCGATGATGT	2892
Qy	2892	AlaThrProProGluIleuAlaAsnGlyValThrGluGlyLeuAspTYrGlyPheMetLys	2911
Db	2893	GCCACCCCGCACAACTGGCCCAAGGGGTACGGAAGCCCTGACCTATGGCTTATGAAG	2952
Qy	2912	GluValThrPheHisCysHisGluGlyTYrIleLeuHisGlyAlaProLysLeuThrCys	2931
Db	2953	GAACTATACATTCCTCACTGTCAAGAGGCTATACATTCGACGGTCTCCAAATCACTCGT	3012
Qy	2932	GlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyPro	2951
Db	3013	CAGTCAGATGGCACTGGATGACGAGATTCCTCTCTGTAAACAGTCAACTGTGGACCT	3072
Qy	2952	ProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIle	2971
Db	3073	CCTGAATATCTTGGCCATGCTTCCCTAAAGGTITTCCTTATTCATGGGGGCAATATA	3132
Qy	2972	GlnTYrGlnCysPheProGlyTYrTYrLysLeuHisGlyLysAsnSerSerArgCysLeuSer	2991
Db	3133	CAGATATCAAGCTTCTCTGGTTATTAAGCTCAATGAAATTCATCAAGAGGTGCTCTCC	3192
Qy	2992	AsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThrProVal	3011
Db	3193	AATGGCTCTGGAGTGGACAGCTCACTTCCTCGCTGCTTCGACATGTTCCACACAGTA	3252
Qy	3012	IleGluTYrGlyThrValAsnGlyThrAspPheAspCysGlyValAlaAlaArgIleGln	3031
Db	3253	ATTGAATAATGAACTGTCAATGGGACGAATTTTACCTGTGAAAGGACGCCGATTCAG	3312
Qy	3032	CysPheIlyGlyPheIlyLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln	3051
Db	3313	TGCTTCAAAGGCTTCAGACTCTTGAACTTCTTAATCACTCTTAAGCCGAGTGGCCAG	3372
Qy	3052	TrpSerSerGlyPheProHisCysGluHisIsthSerCysGlySerLeuProMetIlePro	3071
Db	3373	TGGAGCTCTGGGTTCCCCCACTGGAACAACCTCTGTGGTCTCTCCAAATGATACCA	3432
Qy	3072	AsnAlaPheIleSerGluThrSerSerTrpLysGluAsnValIleThrTYrSerCysArg	3091
Db	3433	AATCGCTTCACAGTGAAGACCAAGCTCTGGAGGAAATGAAATGAAATCACTTACGCTGACGG	3492
Qy	3092	SerGlyTYrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSer	3111
Db	3493	TCTGGATATGTCATACAAAGCAAGTTCAGATTCGATTCGATACAGAGAAAGGGGTATGGAC	3552
Qy	3112	GlnProTYrProValCysGluProLeuSerCysGlySerProProSerValAlaAsnAla	3131
Db	3553	CAGCTTATCCAGTCTGGAGCCCTTGTCCGTGGGTCCCAACCGTCTGTGGCCAAATCA	3612
Qy	3132	ValAlaThrGlyGluAlaHisIleThrTYrGluSerGluValLysLeuArgCysLeuGluGly	3151

Db	3613	GTGGCACTGAGAGGCGACACACCTTATGAAGTGAAGTGAACCTCAAGATCTCGAAGGT	3672
Oy	3152	TYrThwEaSPThrAaSPThrAaSPThrPheThrCySglnLysaAspGlyArgTrpPhePro	3171
Db	3673	TATACGATGAGTATACAGATACAGATACATTCATCCCTGCAGAAAGATGGTCGTGGTCCCT	3732
Oy	3172	GIuAaGllLeSerCySePProLysLeCySPProLeuProGIuAaMlleThHStleLeu	3191
Db	3733	GAGAAATCTCTCGAGTCTTAATAAAATGCTCTCCCGAANAACATACACATATACTT	3792
Oy	3192	ValHieGIaSPaSPPheSerValaAaArgInValSerValSerCyAlaIeGlyTYr	3211
Db	3793	GTACATGGGGAGCATTTCACTGTGAATAGGCAAGTTCTGTGTACATGTGCAGAAAGGAT	3852
Oy	3212	ThrPheGIuGIValaAsnLleSerValCySglnLeuAspGlyThrTrpGIuProPhe	3231
Db	3853	ACCTTGAAGGGAGTTAACATATACATATATGCTGAGTTATGAACTGGGAACCCATTC	3912
Oy	3232	SeAaPbGluSerCySePProValSerCySgIlySPProGluSerProGIuHieGIyPhe	3251
Db	3913	TCCATGATATCTTGCAAGTCCAGTTCTTGTTGGAAACCTGAAGTCCGAACATGATTT	3972
Oy	3252	ValValGIySerLysTYrThrPheGluSerThrIleIleTYrGlnCySgIuProGIlyTr	3271
Db	3973	GTGGTTGGCAGTAATAATACCTTTGAAGCAAAATTAATTATCAGTGAGACCTGGCTAT	4032
Oy	3272	GIuLeuGIu-----	3274
Db	4033	GAACATAGAAATTTGGCTGTGAATCAATCTGTCTTGACCTTTCTTGAGACAGAAC	4092
Oy	3274	-----	3274
Db	4093	CTCAGCTCAGGTGGAGTGGCTTAGAGGTCCAAATCCAGACCTGTGTTGCCGGGTATCA	4152
Oy	3274	-----	3274
Db	4153	GCAGCAGAGGTCAGAAACAGCGATATTGTGAACCGCAATTCGTCTGCCTGATCATTT	4212
Oy	3274	-----	3274
Db	4213	CCTCGGAAGTTTGTCTCAGAGAAATACCGGCGCATGTGAGGTGTCAGTCCGCCCTCAC	4272
Oy	3275	-----GlyaAsnAaGIuAaArgValCySglnGIuAaAaArgIlnTrpSerGIyGIyAlaA	3292
Db	4273	TGGGGGGGGAAACAGGGAAAGTGTCTGCCAGAGGAACAACAGTGAAGTGAAGGGGTGCA	4332
Oy	3293	IleCyLysGIuThrAArgCySgIuThrProLeuGIuPheLeuAaGIlyValaAaSPlle	3312
Db	4333	ATATCAAAAGAACAGAGTGTGAACCTCACTTGAATTTCTCAATGGAAAGCTGACATT	4392
Oy	3313	GIuAaAaArgThrThGIyProAaValValTYrSerCyAaAaArgIlyTYrSerIeGIu	3332
Db	4393	GAATAACAGACCATCGAACCCACGTGTATATCTCTCAACAGAGGCTTACAGCTTGA	4452
Oy	3333	GIyProSerGIuAlaHieCySgThrGIuAaGIlyHTrpSerHieProValProLeuCyS	3352
Db	4453	GCGCATCTGAGGCACTCGACAGAAATGTGAACCTGAGCCACCCAGTCCCTCTGCG	4512
Oy	3353	LysProAaSPProCyProValProPheValIleProGIuAaMlleuLeuSerGIuLys	3372
Db	4513	AAACCAATTCATGCGCTGTCTTTTGATGATCCGAGAAATGCTCTCTCTGAAAG	4572
Oy	3373	GIuPheTYrValaAspGIuAaAaValSerIleLysCyAaArgGIuGIyPheLeuLeuGIuGIy	3392
Db	4573	GAGTTTATGTGATCAGAATGTGTCCATCAAAATGTAGGAAGTTTCTCTCGAGGGC	4632
Oy	3393	HieGIyIleIleThrCyAaSPProAspGIuThrTrpThrGIuThrSerAlaLysCySgIu	3412
Db	4633	CAAGGCATCATTACTGCAACCCGACAGACGTGCAACAGACAGCAAGCGCAATATGAA	4692
Oy	3412	-----	3412
Db	4693	AGAGATATACACACAGCCCAAGTCCGTGAATTTTCAGCTAGACGATTATGAGATTT	4752

OY	3412	-----	3412	-----	3412
Db	4753	AGAAATGTTATTTTCGCGGAGGGGTTCAGATGCGCACTAGAAA	CAGCTGGCGGA	4812	
OY	3412	-----	3412	-----	3412
Db	4813	GCTTCCACCGAGAAACAAAACAAAAGGAGAAAAAGCAAGTGTACACGCCCCA	4872	4872	
OY	3412	-----	3412	-----	3412
Db	4873	CCTGGAGCCACATGGGGCAGCAGAGCTCCACCCCGCAAGAGAGTGA	CTTCCC	4932	
OY	3412	-----	3412	-----	3412
Db	4933	TGCGGAAATTCGAACTCCAGCCAGGGGTTTATGAACAGACTTGATCTCC	TGAGAT	4992	
OY	3412	-----	3412	-----	3412
Db	5053	CTGCTGCTCTGAGAAATCCAGCAGCGTGTACTAGTGGATTCCCA	CAGCAGATT	5112	
OY	3412	-----	3412	-----	3412
Db	5113	CTGTCTCTCCAAAGGCGAGCTAAGCGCTTTGTAAACGAGTCCCTGATCC	ATGCTC	5172	
OY	3412	-----	3412	-----	3412
Db	5173	CTGATTTGGATAGAACCCCCCAACAAGGGGTACGGATGAGACCCCCCA	CAAGGG	5232	
OY	3412	-----	3412	-----	3412
Db	5233	GTCACGACGACCTTATTAAGAGGTGTCTCTGATCAGATCAGTGCCTCTG	GGGA	5292	
OY	3412	-----	3412	-----	3412
Db	5293	CAGAGCTCCAGAGAAAGAGCAGGACGACCTTGTCTGTCTGACAGCTG	CCCTGGAA	5352	
OY	3412	-----	3412	-----	3412
Db	5353	AAGCACGAATTTGGGACAGAGCTAGATTGATGAATTGAAGAAGTAGG	CTTCAGAAAT	5412	
OY	3412	-----	3412	-----	3412
Db	5413	GGGTATTAATGAATTTGCTGAGGCTAAGGAACATGTTCTAAACCAAT	GCAGAACGCA	5472	
OY	3412	-----	3412	-----	3412
Db	5473	AGAACGAGATAAAAATTAACAGATCCGTTAACAGATTAACAGATT	AGAAAGATG	5532	
OY	3412	-----	3412	-----	3412
Db	5533	TAAATGACTGANTGAGCTGAATAACAACAACGAACTTCACATGCA	CAACAAAC	5592	
OY	3412	-----	3412	-----	3412
Db	5593	AAGGCAACATTCAGTTACGAAATTCAGAGAACCCGATGAATATCTC	CATGAGAGA	5652	
OY	3413	-----	3413	-----	3420
Db	5653	TCAAACCCAGACACATTAATCTCAGGTTCTCCAAAGAAATC-TCATG	TGGTCCACAGCT	5711	
OY	3421	HiSvAIGluhenaIaIeAlaArgGIvAlhIbTyrGlnTyrGlyAspMetIleThrTyr		3440	
Db	5712	CACGTGAAGAAATGCAATTCCTCGAGGGGTACATTATCAATATGAGA	CATGATCACTAC	5771	
OY	3441	SeCybTyrSerGIyTyrMetLeuGIuGIyPheLeuArgSerValCybLeuGIuArgGIy		3460	
Db	5772	TCATGTTACAGTGTATCATGTGTGAGAGGTTCTCGAGAGAGTGTGTGTTAGAAATGGA		5831	

QY 3461 ThrTrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAnglyGly 3480
 |||||
 DB 5832 ACATGACATCACTCTCTATTTCAGAGCTGTCTCGATTTCATCTCGAAATGGGGGC 5891
 QY 3481 IleCysGlnArgProGlnAlaCysSerCysProGlnGlyThrMetGlyArgLeuCysGlu 3500
 |||||
 DB 5892 ATCTGCAACCGCCCAATGCTTCTCTGTCAGAGGCTGAGATGGGCCCTCTGTGAA 5951
 QY 3501 GluProIleCysIleLeuProCysLeuAnglyGlyArgCysValAlaProTYrGlnCys 3520
 |||||
 DB 5952 GAACCATCTGCATTCTTCCCTCTCTGAGACGAGGTGCTGTGGCCCTTAACAGTGT 6011
 QY 3521 AspCysProProGlyTYrTrpThrGlySerArgCysHisThr 3533
 |||||
 DB 6012 GACTGCCCGCTGGCTGAGACGGGGTCTGCTGTCAATACA 6050
 RESULT 9
 AAD58963
 ID AAD58963 standard; DNA; 6153 BP.
 AC AAD58963;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DB Human complement receptor-like DNA (POLY11).
 XX
 KW Human; tumour; inflammatory disorder; vaccine; gene therapy; cytostatic;
 complement receptor-like protein; POLY11; gene; ds.
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FH 1..6153
 CDS /tag= a
 FT /product= "Human POLY11"
 FT /note= "This CDS has translational exceptions"
 PT
 US2003050232-A1.
 XX
 PD 13-MAR-2003.
 PE 19-APR-2001; 2001US-00839446.
 XX
 PR 19-APR-2000; 2000US-0198293P.
 PR 20-APR-2000; 2000US-0198645P.
 PR 25-APR-2000; 2000US-0199476P.
 PR 26-APR-2000; 2000US-0199880P.
 PR 26-APR-2000; 2000US-0200024P.
 PR 26-APR-2000; 2000US-0200025P.
 PR 09-JUN-2000; 2000US-0210809P.
 PR 17-JUL-2000; 2000US-0218591P.
 PR 11-AUG-2000; 2000US-0224610P.
 PR 27-FEB-2001; 2001US-0271814P.
 XX
 PA (TAUP/) TAUPIER R. J.
 PA (PAU/) PADIGARU M.
 PA (SPY/) SPYTEK K. A.
 PA (BURG/) BURGESS C. B.
 PA (VERNET/) VERNET C. A. M.
 PA (FERN/) FERNANDES E. R.
 PA (SHIM/) SHIMKETS R. A.
 PA (LITX/) LIT X.
 PA (MAJU/) MAJUMDER K.
 PA (COLM/) COLMAN S. D.
 PA (ZERH/) ZERHUSEN B. D.
 PI Taupier RJ, Padigaru M, Spytek KA, Burgess CB, Vernet CM;
 PI Fernandez BR, Shimkets RA, Liu X, Majumder K, Colman SD;
 PI Zernusen BD;
 DR WPI: 2003-605764/57.
 DR P-PSDB; AAB38612.
 XX

PT New POLYX nucleic acid, useful for preparing a composition for treating
 or preventing e.g., tumor or inflammatory disorder.
 XX
 PS Claim 8; Page 28-31; 75pp; English.
 XX
 CC The invention relates to new POLYX nucleic acid useful for preparing a
 composition for treating or preventing tumor or inflammatory disorder.
 CC The invention is useful as vaccine and in gene therapy. The nucleic acid
 CC is useful for preparing a composition for treating or preventing e.g.,
 CC tumour or inflammatory disorder. The present sequence is human complement
 receptor-like DNA (POLY11)
 CC
 XX
 SQ Sequence 6153 BP; 1610 A; 1486 C; 1559 G; 1497 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 6153
 Score: 8391.00 Matches: 1530
 Percent Similarity: 78.39% Conservative: 1
 Best Local Similarity: 78.34% Mismatches: 10
 Query Match: 42.01% Indels: 412
 DB: 9 Gaps: 3
 US-09-977-053-4 (1-3571) x AAD58963 (1-6153)
 QY 1993 TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyTyrTrpSerArgSer 2012
 |||||
 DB 193 TATACCTTCTGCTGCTCTGACCATTTGAATGCTGACCGCAAGGAGTGAAGAAT 252
 QY 2013 AspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAlaPheIleAlaSerPro 2032
 |||||
 DB 253 GACCAAGATCCCTGCTGCTCTCTGATGATGAGCACCATTTGAGACAGCCCTCTTCA 312
 QY 2033 GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSer 2052
 |||||
 DB 313 GAGACTGCCCATCGGCTCTTGGAGACATTCATCTGATCTGATGATGAGTCAAGC 372
 QY 2053 LeuAlaAspAsnSerGlnLeuLeuGluCysAsnAlaGlnGlyTyrTrpValProProGluGly 2072
 |||||
 DB 373 CTAGACAGCAATTCCTCCAGCTTCTCTGCAATGCCAGGCAAGTGTACCCCGAAGGT 432
 QY 2073 GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer 2092
 |||||
 DB 433 CAGACATGCCCTGCTGATAGCTCATTTCTGTCAAAACCTCCATGCTTCTATAGC 492
 QY 2093 IleLeuGlnSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys 2112
 |||||
 DB 493 ATCTGGATCTGTGAGCAAGCAAAATTTGACGCTGCTCAGTTGTGAGCTTTAAATGC 552
 QY 2113 MetGluGlyPheValLeuAlaThrSerAlaLysIleGluCysMetLysGlyGlyGlnTrp 2132
 |||||
 DB 553 ATGAAAGCTTTGTACTGACACCTCAGCAAGATTAATGATGAGAGGTGGGCAAGTG 612
 QY 2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyLysProProSerIle 2152
 |||||
 DB 613 AACCTTCCCCCATGTCATCCAGTGCATCTCTGCGGTGTGAGAGCCACCAAGATC 672
 QY 2153 MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
 |||||
 DB 673 ATGAATGGCTATGCAAGTGAATCAAACTACAGTTTGTGAGCCATGAGGCTTACAGCTGC 732
 QY 2173 AsnLysGlyPheTyrIleLysGlyLysValLysSerThrCysGluAlaThrGlyGlnTrp 2192
 |||||
 DB 733 AACCAAGGGGTTTCAACATCAAGGAGAAAGAGCACTTCAGAGCCAGGCAAGTG 792
 QY 2193 SerSerProIleProThrCysHisProValSerCysGlyGluLysProProLysValGluAsn 2212
 |||||
 DB 793 AGTAGTCTTATACGAGCTGCCACCGGTATCTTGTGTGAACACCACTAAGGTTGAAGAT 852
 QY 2213 GlyPheLeuGlnHisThrThrGlyArgGlyPheGlnSerGluValAlaGlyTyrGlnCysAsn 2232
 |||||
 DB 853 GGCTTCTGAGAGCTACCACTGCGAGAGATCTTGTGAGGTGAGAGTACAGTGTAC 912
 QY 2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnIleAsnArgHisTrpHis 2252

```
Db 913 CCGGGCTAATAGTCAGTCGGAAGATCCGTGATTTGTCTGCCAAGCAATGCGCATGCGCAC 972
Qy 2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProL1eGlnAsn 2272
Db 973 AGGAAATCCCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
Qy 2273 GlyPheMetLysGlyLysAsnPheGluValGlySerLysValGlnPhePheCysAsnGlu 2292
Db 1033 GGGCTCAATGAAGAGAAACTTGAAGTACGATGATGATGATGATGATGATGATGATGATGAT 1092
Qy 2293 Gly---TyrGluLeuValGlyAspSerSerTyrPheCysGlnLysSerGlyLysTyrAsn 2311
Db 1093 GGGGTATNAGAGCTTGTGTGGGAGACGTTCTTGACATGTGACAAATCTGGCAAAATGGAAAT 1152
Qy 2312 LysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGlnAsn 2331
Db 1153 AAGAAAGTCAATCCAAAGTGCATGCTGCGCAAGTGCACAGCCGCCCTCTTGGAAGAAC 1212
Qy 2332 GlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLysGlu 2351
Db 1213 CAGCTAGTATTAAAGAGATGACCAAGAGTACAGATGATGATGATGATGATGATGATGATGAT 1272
Qy 2352 GlyIleValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyrAsnAsp 2371
Db 1273 AGGCATGCTCTGCAAGGCCCTCTGTCTGAAATGCTTGACATCCGACGAAATGAAATGAC 1332
Qy 2372 SerPheProValCysLysIleValLeuCysThrProProProLeuL1eSerPheGlyVal 2391
Db 1333 TCTTCCCTGTTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1392
Qy 2392 ProIleProSerSerAlaLeuIlePheGlySerThrValLysTyrSerCysValGlyVal 2411
Db 1393 CCAATTCCTTCTTGTCTGCTCTTCAATTTGGAAGTACTGTCAGAGTATCTTGATGATGATG 1452
Qy 2412 PhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTyrSerSerPro 2431
Db 1453 TTTTTCCTAAGAGAAATTCACACCCCTGCGCAACTGAGTGCATGAGCTCTCA 1512
Qy 2432 LeuProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIleIle 2451
Db 1513 CTGCAGAAATGTGTCCAGTGAATGTCCCAACCTGAGAAATCCCAATGGAATTCAT 1572
Qy 2452 AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGlu 2471
Db 1573 GATGTGAGAGCTGTGCTATCTACGACAGCTCTCTATCTGACACCGAGCTTTGAA 1632
Qy 2472 LeuValGlyAsnThrThrThrLeuCysGlyLysGluAsnGlyIleTyrLeuGlyValPro 2491
Db 1633 TTCTGGGGAATATCTACACCCCTTTGTGAGAAATGTGCTGCTGAGAGAAACCA 1692
Qy 2492 ThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyr 2511
Db 1693 ACAGTAAAGCATTTGAGTGTGCTGAAACCAAGAGATTTGAAATGGCAAAATCTCTTAC 1752
Qy 2512 ThrAspLeuIleSyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGlu 2531
Db 1753 AGGAGCTTACACTATGACAGACCTTACTACTCTTGCAACCGAGCTTTGGGCTCCAA 1812
Qy 2532 GlyProSerAlaLeuThrCysLeuGluThrGlyAspTyrAspValAspAlaProSerCys 2551
Db 1813 GGTCCGAGTGCCTTGACCTGTTAGACAGGTGATGGGATGTAGATGCCCATCTTGCC 1872
Qy 2552 AsnAlaIleIleSyrCysAspSerProGlnProIleGluAsnGlyPheValGlnGlyValAsp 2571
Db 1873 AATGCATCCACTGATGATTTCCCAACACCATTTGAAATGTTTGTGAAAGTGCAGAT 1932
Qy 2572 TyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyIle 2591
Db 1933 TACAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
Qy 2592 AlaMetGlnThrCysGluGluSerGlyTyrPheSerSerIleProThrCysMetProIle 2611
Db 1993 GCCATGACAGCTGTGAAAGTCAAGATGCTCAAGTTCCATCCCAATGATATGCCAAATA 2052
Qy 2612 AspCysGlyLeuProProIleIleAspPheGlyAspCysThrLysLeuLysAspAspGln 2631
Db 2053 GACTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2112
Qy 2632 GlyTyrPheGluGlnGluAspAspMetGluValProTyrValIleThrProIleProPro 2651
Db 2113 GATATATTTGAGCAAGAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2172
Qy 2652 TyrIleLeuGlyAlaValAlaIleThrThrGluAsnThrLysGluSerProAlaThrIle 2671
Db 2173 TATCATTTGGAGAGAGTGGCTTAAACCTGGGAAATACAAAGATGATCTCTGCTACAT 2232
Qy 2672 SerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeu 2691
Db 2233 TCAATCAACTTCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
Qy 2692 LeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTyrAsnGlySerAlaProSer 2711
Db 2293 CTGGGGAACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
Qy 2712 CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
Db 2353 TGCATTTCAATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2412
Qy 2732 GluThrSerMetGlySerAlaValAlaGlnTyrSerCysLysProGlyLysIleLeuAlaGly 2751
Db 2413 GAGACTAGCATGGAGAGTGTGCTGACAGTATGATGATGATGATGATGATGATGATGATGAT 2472
Qy 2752 SerAspLeuArgLeuCysLeuGluAsnArgLysTyrSerGlyAlaSerProArgCysGlu 2771
Db 2473 TCTGACTTAAAGCTTGTGTAGAGAAATGAAAGTGAAGTGTGCTCTCCACGCTGTGA 2532
Qy 2772 AlaIleSerCysLysLysProAsnProValIleLeuAsnGlySerIleLysGlySerAsnTyr 2791
Db 2533 GCCATTTATCAAAAGCCAAATCCAGTCAATGATGATGATGATGATGATGATGATGATGATGAT 2592
Qy 2792 ThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGlu 2811
Db 2593 ACATACCTGACAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
Qy 2812 ArgThrThrCysGlnAspArgLysAsnTyrAspGluAspGluProIleCysIleProVal 2831
Db 2653 AGGAGACATGCCAGATGACAAAGCTGGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2712
Qy 2832 AspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGlyIleTyrPhe 2851
Db 2713 GACTGAGTCAACCCCACTGACCAATGCGCAATGCGAGAGAGAGAGAGATGACATTC 2772
Qy 2852 GlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyValArgSerArg 2871
Db 2773 CAAGAAAGATGAAATATACATTCGCAATGAGGATCTTGAGAGAGAGAGAGAGAGAGAGAG 2832
Qy 2872 ValCysLeuAlaAsnGlySerTyrSerGlyValIleThrProAspCysValProValArgCys 2891
Db 2833 GTTGTCTTTCGCAATGAAAGTGTGAGTGAAGTCACTCCGACTGTGTGCTGTGAGAT 2892
Qy 2892 AlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLys 2911
Db 2893 GCCACCCCGCAACATGCGCAATGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2952
Qy 2912 GluValThrPheIleCysIleGluGlyTyrIleLeuIleGlyValProLysLeuThrCys 2931
Db 2953 GAAATGATTCATCCACTGTCAAGAGGCTTACATCTTGCAAGGTGCTCCAAATCTCACTGT 3012
Qy 2932 GlnSerAspGlyAsnTyrAspAlaGluIleProLeuCysLysProValAsnCysGlyPro 2951
Db 3013 CAGTGAATGCGCACTGAGATGAGAGATGCTCTCTGTAAACAGTCAACTGTGACT 3072
Qy 2952 ProGluAspLeuAlaIleGlyPheProAsnGlyPheSerPheIleIleGlyValIle 2971
Db 3073 CCTGAAGATCTTGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3132
```

QY 2972 GlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAsnSerSerArgCysLeuSer 2991
 DB 3133 CAGATCAGAGGCTTCTCGGTATATAGCTCATGGAATTCATGAAAGAGGCTCTCC 3192
 QY 2992 AsnGlySerTTPSerGlySerSerProSerCysLeuProCysArgCysSerThrProVal 3011
 DB 3193 AATGGCTCCCGAGATGCGACGCTCACTTCGCTGCTTCGAGATGTTCCACACCAATA 3252
 QY 3012 ILGluTyrGlyThrValAsnGlyTThrAspPheAspCysGlyLysAlaAlaArgILEGln 3031
 DB 3253 AATGAAATATGAACTGTCATATGGAACAGATTTTGACTGTGAAGGCGACCCGGAATCAG 3312
 QY 3032 CysPheLysGlyPheLysLeuGlyLysSerGlyLysSerGlyLysAlaAspGlyGln 3051
 DB 3313 TGGCTTCAAAAGGCTTCAAGCTCTTAGGACTTCTTGAAATCACCTGTGAAGCGGATGGCAG 3372
 QY 3052 TTPSerSerGlyPheProHisCysGlyLysHisSerCysGlyLysSerLeuProMetILEPro 3071
 DB 3373 TGGAGCTCTGGGTTCCCGCCACCTGTGAACACACTTCTGTGGTTCTCTTCCAAATGATACCA 3432
 QY 3072 AsnAlaPheILESerGlyThrSerSerTTPLysGlyLysAsnValILEThrTyrSerCysArg 3091
 DB 3433 AATGCTTCAATCACTGAGACACCACTCTTGAGAGAAATGTGATTAACGCTGACAG 3492
 QY 3092 SerGlyTyrValILEGlnGlySerSerAspLeuILECysThrGlyLysGlyValTTPSer 3111
 DB 3493 TCTGGATATGTCATACAGGACGATTCGATTTGTATACAGAGAAAGGGGTATGAGC 3552
 QY 3112 GlnProTyrProValCysGlyLysProLeuSerCysGlySerProProSerValAlaAsnAla 3131
 DB 3553 CAGCTTATCCAGCTGTGAGCCCTTGTCTCTGGGTCCACCGTGTGGCCAAATGCA 3612
 QY 3132 ValAlaThrGlyGlyLysHisSerTyrGlyLysSerGlyLysLysLeuArgCysLeuGly 3151
 DB 3613 GTGGCAATCGAGAGGACACCACTTGAAGTGAAGTGAAGTCAAGATGCTGAGAGGT 3672
 QY 3152 TyrThrAspThrAspThrAspThrPheThrCysGlyLysAspGlyArgTTPPhePro 3171
 DB 3673 TATACGATGATACAGATACAGATACATTCACCTGTGAGAAAGATGGCTGGTCCCT 3732
 QY 3172 GlnArgILESerCysSerProLysGlyCysProLeuProGlyLysAsnILEHisILELeu 3191
 DB 3733 GAGAGATCTCTGCACTCTTAAATAATGTCTCTCCGAAACATTAACATATATT 3792
 QY 3192 ValHisGlyAspPheSerValAsnArgGlnValSerValSerCysAlaGlyGlyTyr 3211
 DB 3793 GTACATGGGACCATTTCACTGTGAATATGAGCAAGTTCTGTGTCAATGCCAAGGGTAT 3852
 QY 3212 ThrPheGlnGlyValAsnILESerValCysGlyLysLeuAspGlyThrTTPGlyProPhe 3231
 DB 3853 ACCCTTGAAGGAGTTACATATCAGTATGTCAGCTTATGGAACCTGGAGACCAACATTC 3912
 QY 3232 SerAspGlySerCysSerProValSerCysGlyLysProGlyLysProGlyLysILEPhe 3251
 DB 3913 TCCGATTAATCTTGCACTGCTTCTGTGGAAACCTGAAATGCCAAGACATGAGATT 3972
 QY 3252 ValValGlySerLysTyrThrPheGlnSerThrILETyrGlnCysGlyProGlyTyr 3271
 DB 3973 GTGGTTGGCATTAATCACTTTGAAGCACATTTATTTATCACTGTGAGCCGTGGCAT 4032
 QY 3272 GlnLeuGln----- 3274
 DB 4033 GAACCTAGAAATTTGGCTGTAATCATCTGCTCTGGAATTTCTTGTTGACAGAGAC 4092
 QY 3274 ----- 3274
 DB 4093 CTCAGCTGCAAGTGGAGTTGGCTAGAGGTCCAATCCAGACCTGTTGGCTGGGTATCA 4152
 QY 3274 ----- 3274
 DB 4153 GCAGCAGAGGCTGACAGAAAGCGGATATTGGTGAACCGCAAAATGCTGCTCCATCATTT 4212

QY 3274 ----- 3274
 DB 4213 CCTTGAAGATTTGTCTCAGAGGAATACCCGGCCATGTAGGTGTCACTCGCCCTAC 4272
 QY 3275 -----GlyAsnArgGlyLysValCysGlnGlyLysAsnArgGlnTTPSerGlyValValAla 3292
 DB 4273 TGGGGGGGAAACAGGAACGATGTCTGCAAGAGAAACACAGTGAAGTGAAGGGGTGGCA 4332
 QY 3293 ILecLysGlyThrArgCysGlyLysThrProLeuGlnPheLeuAsnGlyLysAlaAspILE 3312
 DB 4333 AATGCAAAAGACACAGGTGTGAATCTCACTTGAATTTCTCAATGGGAATCTGCAAT 4392
 QY 3313 GlnAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGln 3332
 DB 4393 GAAACAGAGACACTGAGACCAAGTGTATATCTTCCACAGAGGCTACAGTCTTGA 4452
 QY 3333 GlyProSerGlyLysHisCysThrGlyLysAsnGlyThrTTPSerHisProValProLeuCys 3352
 DB 4453 GGGCATTTGAGGCACTGTGACAGAAATGGAACCTGAGCCACCCAGTCCCTCTTGC 4512
 QY 3353 LysProAsnProCysProValProPheValILEProGlyLysAsnAlaLeuLeuSerGlyLys 3372
 DB 4513 AAACCAATTCATGCTCTGCTCTTTGTGATTCGGAATGCTCTGCTGTCTGAAG 4572
 QY 3373 GluPheTyrValAspGlnAsnValSerILECysArgGlyLysPheLeuLeuGlnGly 3392
 DB 4573 GAGTTTATGTTGATCAGAAATGTGTCCATCAATGTAGGAAAGTTTCTGCTGACGGCC 4632
 QY 3393 HisGlyILEThrCysAsnProAspGlyLysThrTTPThrGlnThrSerAlaLysCysGly 3412
 DB 4633 CAGGCATCATTTACTGTGACCCCGACAGAGGTGTGACAGACAGAGCCCAATGTGA 4692
 QY 3413 Lys----- 3413
 DB 4693 AGAAGATATACAAAGCCCAAGTCCCTGAATTTTACGTAAGCAGCTTATTCAGTATT 4752
 QY 3413 ----- 3413
 DB 4753 AGAATGTTATTTGCGGGAGGGGTTCAGATGCGCACTAGAAACAGCTGTGCGGA 4812
 QY 3413 ----- 3413
 DB 4813 GCTTCCACCGAGAAACAAACAAACGAGAAAAAGAGGTGTACACAGGCCCA 4872
 QY 3413 ----- 3413
 DB 4873 CCTGGAGCCATGGGGCAGAGAGCTCCACCCCGCAAGAGGTGGAAGCTCC 4932
 QY 3413 ----- 3413
 DB 4933 TGGGGAATTTACAGCACTCCAGCCAGGGTTATGAAACAGACTGTATCTCTGAGA 4992
 QY 3413 ----- 3413
 DB 4993 TGAAGCCCTGGGCTCATATGTGCAATGTCTTCCACAGATCAGAGGCTTATGCTTCC 5052
 QY 3413 ----- 3413
 DB 5053 CCTGTGCTGTGAGGAATCCAGGAGGCTGAGACTAGTGGATTCACACAGACAGTTT 5112
 QY 3413 ----- 3413
 DB 5113 ACCGTCTTGCAGAGGGGACGTAGAGGCTTTTGAAGCAGATCCCTGATCCATGCT 5172
 QY 3413 ----- 3413
 DB 5173 CCTGATTTGGATGAGACCCCGCCACACAGAGGCTCACAGATGAGACCCCGCCACACAG 5232
 QY 3413 ----- 3413
 DB 5233 GGTACACAGACCTTATACAGAGGTGTTCTGTAGCATCAGGTCAAGTCCCTGTGG 5292
 QY 3413 ----- 3413

```

Db 5293 ACAGAGTCCGAGAGAAAGAGAGGAGGAGCATCTTGCTGTTCTGAGGTCGGCTGGA 5352
QY 3413 ----- 3413
Db 5353 AAGGACAGAAATTTGGGAGAGGCTAGATTTGATTTGAAAGAGAGCTTACAGAAAG 5412
QY 3413 ----- 3413
Db 5413 TGGGTAATATATGAAGTTGCTGAGCTAAAGAAACATGTTCTTAAACCATGCAAGAGGCC 5472
QY 3413 ----- 3413
Db 5533 GTAATGACCTGATGAGCTGAAAAACAGACAGAGAACTTCAATGCAACAGACAAA 5592
QY 3413 ----- 3413
Db 5593 CAAGGCCAATTCAGTTCCAGGAAATCCAGAGAACCCAGTAAGTACTCCATGAGAG 5652
QY 3414 ----- 3414
Db 5653 ATCAACCCAGAGACATATATCTCAGGTTCTCCAGAAATCTCATGTGTCACAGGCT 5712
QY 3421 HisValGluAspAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyr 3440
Db 5713 CACGTAGAAATATGCAATTTCTCGAGGCTGACATTTATCAATATGAGACATGATCACCTAC 5772
QY 3441 SerCysTyrSerGlyTyrMetLeuGlyPheLeuArgSerValCysLeuGluAsnGly 3460
Db 5773 TATGTTACAGTGGATACATGTTGAGGGGTTCTCTAGAGAGTTGTTTGAAGAAATGA 5832
QY 3461 ThrTyrThrSerProGlyCysArgAlaValCysArgPheProCysGluAsnGly 3480
Db 5833 ACATGACATACACTCTCTATTTTCAGAGGCTGTGTGATTTCCATGTCAGATGAGGAGG 5892
QY 3481 IleCysGluArgProAspAlaCysSerCysProGluGlyTyrMetGlyValGluCysGlu 3500
Db 5893 ATCTGCAACGCGCAATGCTTCTCTCTCCAGAGGCTGAGATGGGCGCCCTCTGTGA 5952
QY 3501 GluProIleCysIleLeuProCysLeuAsnGlyValArgCysValAlaProTyrGlnCys 3520
Db 5953 GAACCAATCTGCAATTTCTCCCTCTCTGAAAGAGGCTGCTGTGGCCCTTACAGGT 6012
QY 3521 AspCysProProGlyTyrThrGlySerArgCysHisThr 3533
Db 6013 GACTGCCGCGCTGGCTGAGAGGGGCTCGCTGTCAATACA 6051

```

RESULT 10

AAH16667 standard; cDNA; 5124 BP.

AAH16667;

26-JUN-2001 (first entry)

Human cDNA sequence SEQ ID NO:15810.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000BP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX
 XX (HELI-) HELIX RES INST.
 XX Oca T, Isogai T, Mishioka T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR
 XX
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.

Claim 8; SEQ ID NO 15810; 2537bp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH16628 and
 CC AAH16633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

Sequence 5124 BP; 1462 A; 1106 C; 1178 G; 1378 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 5124
 Score: 7537.00 Matches: 1347
 Percent Similarity: 99.63% Conservative: 1
 Best Local Similarity: 99.56% Mismatches: 5
 Query Match: 37.74% Indels: 0
 DB: 4 Gaps: 0

US-09-977-053-4 (1-3571) x AAH16667 (1-5124)

```

QY 195 ThrAspGlyTyrSerAsnGlyValAspProArgProIleAlaIleSerLeuAlaAspSer 214
Db 2 ACTGATGATATTCCAATGCGGAGACCTTAGACCAATTCAGAGTCACTGCGAGATTCA 61
QY 215 GlyValGluIlePheThrPheGlyIleTyrGluGlyValMetLeuArgGluLeuAspMet 234
Db 62 GAGGTGAGATCTTCACTTTTGCAATGCGAGAGAACATTCAGAGCTGAAAGACATG 121
QY 235 AlaSerThrProArgGluIleValCysTyrLeuLeuHisSerPheGluGluPheGluAla 254
Db 122 GCTTCACCCCAAGAGAGAGCACTGTACCTGACACAGTTTGAAGATTTAGAGGCT 181
QY 255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerCysSerPheIleGlnAspAspMet 274
Db 182 TTAGCTGCGCGGAGATTGCAATGAGATCTTCTGAGGTTTATTCAAGATGATATG 241
QY 275 ValHisCysSerTyrLeuCysAspGluGlyValAspCysCysAspArgMetGlySerCys 294
Db 242 GTCCACTGCTCATCTTCTTTGATGAAAGGCAAGCACTGCTGTACCGAATGGAGAGCTGC 301

```

QY 295 LysCysGlyThrHisThrGlyHisPheGluCysIleCysGlyValGlyTyrTyrGlyVal 314
| | | | |
DB 302 AATATGGGACACACACAGGCACTTTTGAAGTCATCTGTGAAAAGGGGATATTACGGGAAA 361
QY 315 GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro 334
| | | | |
DB 362 GGTCTGACATATGAAAGACAGCTTGCCACGGGAACTACAAACCTGAAAGCTCACCA 421
QY 335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGlySer 354
| | | | |
DB 422 GAGGATACGACGATTCATTCATGTCCTGATGAATACACCTCTCCACCTGGAAAC 481
QY 355 ThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyValInThCysGlu 374
| | | | |
DB 482 ACATCCCTCAAGACCTGTCTGACGAGAGGATACAGGGCATCTGGCACAACCTGGAA 541
QY 375 LeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCys 394
| | | | |
DB 542 CTGTCTCACAGCCCTGCTGACAGCTCCGAAATGGTTACTTTTCCAAAACACTTGC 601
QY 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
| | | | |
DB 602 AACAACCACTTCATGACGACCTGTGGGGTCCGATGTCACTTGATTTGATCTGTGGGA 661
QY 415 SerSerIleIleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArg 434
| | | | |
DB 662 AGCAGCATCATCTTATGCTTACCCCAATGGTTTGTGGTCCGGTTTAAAGAGCTACTGCGA 721
QY 435 ValArgThrCysProHisIleLeuArgGlnProLysHisGlyHisIleSerCysSerThrArg 454
| | | | |
DB 722 GTTAAAGAACAGTCTCACTCCGACCGAACAATGGCCACATCAGCTGTCTTACAGG 781
QY 455 GluMetLeuTyrIleSerThrThrCysIleValAlaCysAspGluGlyTyrArgLeuGluGly 474
| | | | |
DB 782 GAATGTTATATATAGACACACATGTTGGTTGCTGTGATGAAGGGTACAGCTGAAGGC 841
QY 475 SerAspLysLeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysVal 494
| | | | |
DB 842 AGGATATAGCTTACTTGTCAAGAAACAGCCAGTGGAGTGGCCAGAACCCCGGTGTGTG 901
QY 495 GluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnGly 514
| | | | |
DB 902 GACGCGACAGTCTCCACTTTCAGATGCCAAGATGTCATCATATCCCCCAACACTGT 961
QY 515 GlyLeuGlnProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIle 534
| | | | |
DB 962 GCGAAGACGACGACCAATTTGGACGATCTGATGTAAGTTCCCGCAAGGGTTCACTT 1021
QY 535 LeuSerGlyValIlyGluMetLeuArgCysThrThrSerGlyValTyrPheAsnValGlyVal 554
| | | | |
DB 1022 TTATCTGGAGTCAAGAAATGCTGAGATGTACACTTCTGGAAAATGGAATGGGAGTT 1081
QY 555 GlnAlaAlaValCysLysAspValGlnAlaProGlnIleAsnGlyProLysAspIleGlu 574
| | | | |
DB 1082 CAGGCACTGTGTGTAAAGACGTGAGGCTCTCAATCAACTGTCTTAAGACATAGAG 1141
QY 575 AlaLysThrLeuGlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLys 594
| | | | |
DB 1142 GCTAAGACTCTGGAAACGAAAGATTGCCAATGTACTCGACAGATTCCAAACAGCTAAA 1201
QY 595 AspAsnSerGlyValLysValSerValHisValHisProAlaPheThrProProTyrLeu 614
| | | | |
DB 1202 GACAACTCTGGTAAAGGTGTGCTGACGTCACGTCATCAAGCTTTCACCCCACTTACCTT 1261
QY 615 PheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyLysGlnAla 634
| | | | |
DB 1262 TTCCCAATTCGAGATGTGTCTATCTGATACAGGCACTGACCTATCCGGAACAAGCC 1321
QY 635 SerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTyrCysArg 654
| | | | |
DB 1322 AGCTGACATTTTCCATATCAAGGTATATGATGCAAGAACACCTGTCAATAGCTGTGGAGA 1381
QY 655 SerProProProValGlnValSerGlyValHisAlaIleSerTyrAspGluProGln 674

DB 1382 TCTCACCTCCGTCAGGTCTCGAGAAAGTTACATGCCAGCTGGAGATGACCTCAG 1441
QY 675 PheSerAspAsnSerGlyValGluLeuValIleThrArgSerHisThrGlnGlyAspLeu 694
| | | | |
DB 1442 TTCTCAGAACACTCAGGGGCTGAAATTTGGTCACTTACCAAGATCATACCAAGGACCTT 1501
QY 695 PheProGlnGlyValThrIleValIleGlnTyrThrAlaThrAspProSerGlyLysAsnArg 714
| | | | |
DB 1502 TTCCCTCAAGGGGAGACTATATGACAGTATACGGCACCTGACCCCTCAGGCAATTAACAG 1561
QY 715 ThrCysAspIleHisIleValIleLysGlySerProCysGlnIleProPheThrProVal 734
| | | | |
DB 1562 ACATGTATATCCAAATATGTATAAAGGTTCTCCTGTGAATTCATTCACACTGGA 1621
QY 735 AsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnGlyThrLeuThrCysLeu 754
| | | | |
DB 1622 AATGGGATTTTATATATGACCTCCAGTATATCTGGAGTCAACTGTATCACTTACTTGCCTT 1681
QY 755 GluGlyTyrAspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGly 774
| | | | |
DB 1682 GAGGGCTATGATTTACAGAAAGGCTTACTGACAAATATATGTGTATGTCTTATCAAGATGCG 1741
QY 775 ValTyrLysProThrTyrThrThrThrGluThrProAspCysAlaLysLysArgPheAlaAsn 794
| | | | |
DB 1742 GTCTGAAACCAACATATATACCACTGAATGGCCAGACTGTGCGAAAACGTTTGCACAC 1801
QY 795 HisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeu 814
| | | | |
DB 1802 CACGGGTTCAAGTCTTTGAGATGTTCTTCAAAAGCAGCTCGTGTGATGACACAGATCTG 1861
QY 815 MetLysLysPheSerGlyAlaPheGluThrThrLeuGlyLysMetValProSerPheCys 834
| | | | |
DB 1862 ATGAAGAAAGTTTCTGAAGCACTTGACAGCACTCCGGGAAAATGCTCCCACTATTGT 1921
QY 835 SerAspAlaGluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeu 854
| | | | |
DB 1922 AGTATGACAGAGACATTTGACTGAGACTGGAGAGAACCTGACCAAAAATATATGCTTA 1981
QY 855 GluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrPglValAla 874
| | | | |
DB 1982 GAATATATATATATGCTATGAAATGCGCTTTCATTTGACCAAGGCTGGGGTGCAGCT 2041
QY 875 AsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGluInThrAlaThrSer 894
| | | | |
DB 2042 AATAGGCTGATTACTCTTAAAGATGACTTCTGACACTGTGCAAGAAACAGCCCAAGC 2101
QY 895 IleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIle 914
| | | | |
DB 2102 ATCCGCAATGCCAAGTCTCTCAGGATTTAAAGAGTGGCCCATTTTCTGACTATTAATTT 2161
QY 915 LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluAsnAspThrLeu 934
| | | | |
DB 2162 AAGTTAATTTTAAATCATCAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCTT 2221
QY 935 GluTyrPgluAsnGlnGlnArgLeuLeuGlnThrLeuGlnThrIleThrAsnLysLeuLys 954
| | | | |
DB 2222 GAATGGGAAATTCAGCAACGACTCTTCAGCAATTTGAAACATATCACAATTAACGTGAAA 2281
QY 955 ArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAla 974
| | | | |
DB 2282 AGGACTCTCAACAAAGACCCCAATGATTTCTTTCAGCTTGCAATCAAGAAATCTTATACC 2341
QY 975 AspSerAsnSerLeuGlnThrLysValAspProPheCysArgProGlySerValLeu 994
| | | | |
DB 2342 GACAGCAATTCATTAGGAACAAAAAGGCTTCCCTCTCAGACCGAGCTCAGTGTG 2401
QY 995 ArgGlyArgMetCysValAsnGlyProLeuGlyThrTyrTyrAsnLeuGlnHisPheThr 1014
| | | | |
DB 2402 AGAGGGGTATGTGTCAATTTGCTTGGAACTATTAATCTGAAACATTTACCC 2461
QY 1015 CysGluSerCysArgIleGlySerTyrGlnAspGluGluGlnLeuGluCysLysLeu 1034
| | | | |

Db 2462 TGTGAAGCTGCCGATCCGATTCATCAAGATGAAGAGGCACTGATCCAGCTT 2521
 Qy 1035 CysProserGlyMetThrGluTyrIleHisSerArgAsnIleSerAspCysValAla 1054
 Db 2522 TCCCTCTCGGAGATGACACGGAAATATTCATCAAGAAACATCTCTGATGTAAAGCT 2581
 Qy 1055 GlnCysValGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeu 1074
 Db 2582 CAGGTAAACAGAGGACCTACTCATGACAGCTTGAGACTTGATGATGTGTCCACTG 2641
 Qy 1075 GlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
 Db 2642 GGCACCTATCAGCAAAATTTGGTCCCGAGCTGCTCTCGTCCAGAAAACACTCA 2701
 Qy 1095 ThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLys 1114
 Db 2702 ACTGTAAAAAGAGACCGTGAACATTTCTGACATGAGAGTCTCTGTCCAGAAAGAAA 2761
 Qy 1115 PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsn 1134
 Db 2762 TTCTCCGCTTCTGGGTTAATGCTCTGACCCATGCTCTGACTATTAACAACTAAT 2821
 Qy 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySer 1154
 Db 2822 GAGGGAGAGGCTTCTGCTGCTGCTGCTCTCTTTATGAACTACCCCATTCGCTGCTTCC 2881
 Qy 1155 ArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerVal 1174
 Db 2882 AATTCATCAACAGAAAGTTCAAGTTTAACTTCACTTCACGCGGAGAGAAAGTG 2941
 Qy 1175 ValProProAlaSerLeuGlyHisIleLysValArgHisGluIleSerSerGlnValPhe 1194
 Db 2942 GTGCCCCCTGCTCTCTGACATATTAAAGAGCATGAATACAGCATCGATTTTC 3001
 Qy 1195 HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGluLeuGlyValGly 1214
 Db 3002 CATGAATGCTTCTTAAACCTTCCCAATATGAGAAACCTGCAACATTTGGCGGTTC 3061
 Qy 1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGlu 1234
 Db 3062 TATGTTGCTCTGTCATCTGATATACAGCTTAAAGTTGTAACAGACATCGATGAG 3121
 Qy 1235 CysSerProLeuProCysLeuAsnAsnGlyValCysLysPheLeuValGlyIlePheIle 1254
 Db 3122 TCAGGCCACCTGCTTGCCTCAACATGAGGTTGTAAGACCTAATGGGGAATTCATT 3181
 Qy 1255 CysGluCysProSerGlyTyrThrGlyGlnArgCysGluAsnIleAsnGluCysSer 1274
 Db 3182 TGTGAGTGCCTCATCAGGTTACACAGGTCACGCGTGTGAAGAAAATTAATGAGTGAAC 3241
 Qy 1275 SerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThr 1294
 Db 3242 TCCAGTCTTGTGTTAATTAAGAAATCTGTGATGAGTGTGCTGCTTTCGTTGACA 3301
 Qy 1295 CysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsn 1314
 Db 3302 TGTGTAAAGGATTTGTAGGCTCGCATTTGGAACAGAACTCATATGATCCAGTAAAC 3361
 Qy 1315 ProCysLeuAsnAsnAlaValCysGluAspGlnValGlyIlePheLeuCysValCysPro 1334
 Db 3362 CCATGCTTAAATTAATCAGTCTGTGAAGACAGAGTTGGGGGATTTCTGTGCAATGGCCA 3421
 Qy 1335 ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCys 1354
 Db 3422 CTTGATTTTGGGTACCCGATGTGAAGAAAGCTGATAGTGTCTCACTGACCCATGC 3481
 Qy 1355 LysAsnGlyAlaThrCysLysAspGlyValAsnSerPheArgCysLeuCysAlaAlaGly 1374
 Db 3482 AAAAATGAGCTACCTGTAAAGACGCTGCAATAGCTTCAGATGCTGTGACAGCTGGC 3541
 Qy 1375 PheThrGlySerHisCysGlyLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsn 1394
 Db 3542 TTCAACGATCACTGTGAATGAACATCAATGAATGTCAGTAAATTCATGATGAAT 3601

Qy 1395 GlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSer 1414
 Db 3602 CAGGCCACCTGTGTGATGAATTAATTAATCAAGTTGTAAGTGAAGCAGGATTTTCA 3661
 Qy 1415 GlyGlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
 Db 3662 GGCAAAAGGTGTAAACAGAACAGTCTACAGGCTTTAACTGTGAATTTGAAAGTTTCGGC 3721
 Qy 1435 IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
 Db 3722 ATCTATGATATATCATGCTAGTTGGCATCTCCCATCTCAATGCTTAACCTGTACC 3781
 Qy 1455 PheTrpMetLysSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAsp 1474
 Db 3782 TTCTGATGAATACCTCTGACGACATGAATGAACACCAATCTCTATGAGCTTAT 3841
 Qy 1475 AsnGlySerAspAsnThrLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
 Db 3842 AACGGACGACAAATACCTTGTCTGACTGATTAATACGCGTGGTTCCTTATGTGAAT 3901
 Qy 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTrpHisAlaIleAla 1514
 Db 3902 GGCAGGGAAGATTAACAACTGCTCCGTGAATGATGAGATGCAATCATTTGCA 3961
 Qy 1515 IleThrTrpThrSerAlaAsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAsp 1534
 Db 3962 ATCATTTGACAACTGTCATGAGCATCTGAAAGTCTATATCAATGGAATTAATTCGAC 4021
 Qy 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
 Db 4022 GTGTGTGCTGCTCTCTGTGTGTGTCCTATCTGCT 4060

RESULT 11
 ABK13101
 ID ABK13101 standard; DNA: 3991 BP.
 AC ABK13101;
 XX
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE Human REPT12 cDNA sequence.
 XX
 XX REPT12; human; anti-inflammatory; cytosolic; immunosuppressive; antiviral;
 KM anti-HIV; anti-arthritic; anticonvulsant; nootropic; neuroprotective;
 KM anti-lymphic; antibody; immunogen; endometriosis;
 KM gastrointestinal disorder; gastritis; oesophageal carcinoma;
 KM Crohn's disease; irritable bowel syndrome; ulcerative colitis;
 KM endocrine disorder; hypochalimus disorder; Kallman's disease;
 KM autoimmune disease; inflammatory disease; infertility; receptor;
 KM acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;
 KM osteoarthritis; diabetes mellitus; multiple sclerosis;
 KM systemic lupus erythematosus; cell proliferative disorder; cancer;
 KM developmental disorder; Duchenne muscular dystrophy;
 KM Becker muscular dystrophy; neurological disorder; epilepsy;
 KM Alzheimer's disease; Huntington's disease; reproductive disorder; ss.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 161..3916
 FT /tag= a
 FT /product= "REPT12 protein"
 XX
 XX
 PN W0200198154-A2.
 XX
 XX 27-DEC-2001.
 PD
 XX
 PF 21-JUN-2001; 2001MO-US019942.
 XX
 PR 21-JUN-2000; 2000US-0214027P.
 PR 25-AUG-2000; 2000US-0228045P.
 PR 12-DEC-2000; 2000US-0255104P.

2684 rCysaenProglYTYrGluLeuLeuGlYAsnProValLeuLeuLeuLeuGlnGluAsnGluYTh 2704
Db 1441 CCGTATCCAGAGATATGAAGCTTCGGGGAACCTGGTGCATCTGCACAGGAATGGAAC 1500
QY 2704 rTTPAsnGlySerAlaProSerCyAlIeserIleGluCyAsnLeuProThrAlaProGl 2724
Db 1501 TTGGATGGCAGGACCACTCCGCAATTCATTTGAATGACCTTCCTCTCTGA 1560
QY 2724 uAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTYrSerCyAlu 2744
Db 1561 AAATGGCTTTTCCGCTTTTACAGAGACTACATGGGAAGCTGTGACGATATAGCTTAA 1620
QY 2744 aProGlyAlaIleLeuAlaGlySerAsnLeuArgLeuCyAluGluAsnArgIleTYrSe 2764
Db 1621 ACCTGGACACATCTAGCAGGCTCTGACTTAAAGCTTTGTCTAGAGAAATGGAAGTGGAG 1680
QY 2764 rGlyAlaSerProArgCyGluAlaIleSerCyAluGlyblybProAsnProValMetAsnG 2784
Db 1681 TGGTGCCTCCCAACGCTGGAAACCATTCATGCAAAAAGCCAAATCCAGTCATGATGG 1740
QY 2784 ySerIlelySGlySerAsnTYrThTYrIleuSerThrLeuTYrTYrGluCyAsnProGl 2804
Db 1741 ATCCATCAAGAGAGACACTACACATACCTGACGACCTTGTACTATGAGTGAACCCCG 1800
QY 2804 YTYrValLeuAsnGlyThrGluArgArgThrCySGlnAspAspLyAsnTTPAspGluAs 2824
Db 1801 ATATGTGCTGAATGGCACTGAGAGAGAAACATGCGAGATGACAAAACCTGGGATGGGA 1860
QY 2824 rGluProIleCyAlIerProValLeuProCySerSerProProValSerAlaAsnGlyGlnVa 2844
Db 1861 TGAGCCCATTTGGATTCCTGTGGACGTGCAATCCCACTCCAGCTCAGCCCATGGCCAGG 1920
QY 2844 lArgGlyAspGluTYrThrPheGlnLySGluIleGluTYrThrCyAsnGlnGluYPhel 2864
Db 1921 GAGAGAGAGAGAGATACATTCCTCAAAAAGATTAATACCTTGCAATGAAGGTTCTT 1980
QY 2864 uLeuGlnGlyAlaArgSerArgValCyAluAlaAsnGlySerTTPSerGlyAlaThrPr 2884
Db 1981 GCTTGAAGGAGCCAGAGATCGGGTTGTCTTCCCAATGGAAGTTGAGTGAACCACTCC 2040
QY 2884 cAspCyValIProValArgCyAlaAlaThrProProGlnLeuAlaAsnGlyValThrGluG 2904
Db 2041 CCACTGTGTGCTGTGACATGTGCAACCCGCAACAATGGCCCAATGGGGTGAAGGAAG 2100
QY 2904 yLeuAspTYrGlyPheMetIleSGluValThrPheIleSGlyAlaGlyTYrIleLeuH 2924
Db 2101 CCTGGAATATGGCTTCAAGAGAAATTAACATTCCTGACAGAGGGCTACATCTTGA 2160
QY 2924 sGlyAlaProLybLeuThrCySGlnSerAspGlyAsnTTPAspAlaGluIleProLeuCy 2944
Db 2161 CCGTGTCCAAAACCTCACTGTGACATGAGCACTGGGATGACAGAGATTCCTCTCTG 2220
QY 2944 blybProValAsnCySGlyProProGlnLeuAlaHISGlyPheProAsnGlyPheSe 2964
Db 2221 TAAACCAAGTCAACTGTGACCTCTGAAGATCTTGGCCATGGTTCCCTTAATGGTTTTC 2280
QY 2964 rPheIleHISGlyGlyAlaIleGlnTYrGlnCySPheProGluTYrTYrLeuHISGlyAs 2984
Db 2281 CTTTATTCATGGGGCCATATACATGATCAAGTCTTCCGAGTTATAGCTCAATGAAA 2340
QY 2984 nSerSerArgArgCyAluSerAsnGlySerTTPSerGlySerSerProSerCyAluPr 3004
Db 2341 TTCATCAAGAGGTGCTCTCCCAATGGCTCTGGAATGGGACGCTCACTTCCGCTGCC 2400
QY 3004 cCyAsnArgCySerThrProValIleGluTYrGlyblybProValAsnGlyblybPheAsp 3024
Db 2401 TTGCAGATGTTCACACCAAGTAATGAATGGAACGTCAAAAGGACAGATTTTACATG 2460
QY 3024 sGlyLybAlaAlaArgGlnGlnCySPheIleSGlyPheLybLeuGlnGlyLeuSerGluI 3044
Db 2461 TCGAAAGGACGCCGAGATTCAGTGTCTCAAGGCTTCAAGCTCTTGAAGCTTTCTGAAT 2520

QY 3044 eThrCySGluAlaAspGlyGlnTTPSerSerGlyPheProHISCySGlnHISThrSerCy 3064
Db 2521 CACCTGTGAAGCCGATGCGACGTGAGACTCTGGGGTCCCCACCTGTGAACACTTCTTG 2580
QY 3064 sGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTTPLySGluAs 3084
Db 2581 TGGTCTCTTCATGATATACAAATGGATGATCAAGAAACCAAGCTCTTGAAGAAA 2640
QY 3084 nValIleThrTYrSerCyAsnSerGlyTYrValIleGlnGlySerSerAspLeuIleCy 3104
Db 2641 TGTGATACCTTACAGGTGACAGGTCTGGATATGTCAATCAAGGACAGTTCAATCTGATTG 2700
QY 3104 sThrGluLybGlyValTTPSerGlnProTYrProValCySGluProLeuSerCySGlySe 3124
Db 2701 TACAGAAAGAGGGGTATGGAGCCACTTATCCAGTGTGAAACCCCTTGTCCCTGTGGTC 2760
QY 3124 rProProSerValAlaAsnAlaValAlaThrGlyGluAlaHISThrTYrGlnSerGluVa 3144
Db 2761 CCAACGCTGTGCGCAATGCAAGTGGCACTGGAAGGACCCACCTATGAAGTGAAGT 2820
QY 3144 llybLeuArgCyAluGlnGluTYrThrMetAspThrAspThrAspThrCySGl 3164
Db 2821 GAATCCAGATGCTGGAAGTTATAGATGATACAGATACAGATTCATCACTGTCA 2880
QY 3164 nlybAspGlybArgTTPPheProGluArgIleSerCySerProLybLybCySPoleuPr 3184
Db 2881 GAAGAATGGTGGCTGTCCCTGAGAGAAATCTCTCAGACTTAAATAATGCTCTCC 2940
QY 3184 oGluAsnIleThrHISIleuValHISGlyAspAspPheSerValAsnArgGlnValSe 3204
Db 2941 GGAATCATTAACATATACTGTTCATGGAGGAGATTCAGTGTGATAGGCAAGTTTC 3000
QY 3204 rValSerCyAlaGlnGluTYrThrPheGlnGlyValAsnIleSerValCySGlnLeuAs 3224
Db 3001 TGTGTATGTGAGAGAGGATATACCTTGAAGGAGTTATACATACGATATGCTGACTGA 3060
QY 3224 rGlyThrTTPGluProProPheSerAspGlyblybSerCySerProValSerCySGlyblybPr 3244
Db 3061 TGGAACTGGAGCCCACTATCTCCAGATGATCTTCAGATGCAAGTTCCTGTGGAAACC 3120
QY 3244 oGluSerProGluHISGlyPheValValGlySerLybTYrThrPheGlnSerThrIleI 3264
Db 3121 TGAAGTCCAGAACATGGAATTTGTGTGGCTGAGTAATACACTTTGAAGCAATAT 3180
QY 3264 eTYrGlnCySGluProGlyTYrGluLeuGlnGlyAsnArgGlyblybValCySGlnGluAs 3284
Db 3181 TTATCAGTGTAGCTCGGCTATGACATGAGGGGAACAAGGAACGTGTGCGCAGAGAA 3240
QY 3284 nArgGlnTTPSerGlyblybValAlaIleCyAluGlyblybThrArgCySGluThrProLeuG 3304
Db 3241 GAGACAGTGAAGTGGAGGGGTGGCAATATGCAAAAGAACCAAGGTGTGAACCTCACTGA 3300
QY 3304 uPheLeuAsnGlyLybAlaAspIleGluAsnArgThrThrGlyProAsnValValTYrSe 3324
Db 3301 ATTTCTCAATGGAAAGCTGACATTGAACAGAAAGACTGAGCCCACTGGTATATTC 3360
QY 3324 rCyAsnArgGlyTYrSerLeuGlnGlyProSerGluAlaHISCySerThrGluAsnGlyTh 3344
Db 3361 CTGCAACAGAGGCTACGCTTGAAGGCACTGAGGCACTGCAAGAAATGGAAC 3420
QY 3344 rTTPSerHISProValProLeuCyAluSerProAsnProCySProValProPheValIlePr 3364
Db 3421 CTGGAGCCCAACCAAGTCCCTCTGCAACCAAAATCCATGCTGTCTTTTGTGATTC 3480
QY 3364 oGluAsnAlaLeuLeuSerGlyblybGlybPheTYrValaAspGlnaAsnValSerIleLybCy 3384
Db 3481 CAGATGCTCTGCTGTCTGAAGAGATTTATGTGATGACAGATGTGCTCAATG 3540
QY 3384 sArgGlnGlyPheLeuLeuGlnGlyblybGlyIleIleThrCyAsnProAspGluThrTr 3404
Db 3541 TACGAAAGGTTTCTGCTCAAGGCAAGGCAATCAATCACTGCAACCCCAAGACAGTGTG 3500
QY 3404 rThrGlnThrSerAlaLybCySGluLybIleSerCySGlyProProValHISblybValu 3424

DB 3601 GACACAGACAGACGCAATGTAATAATCTCATGTGGTCCACGCTCACCTAGAAA 3660
 OY 3424 malialaialaarglvalahistyrghlntyrghlaspmetilethrtyrsercyserylse 3444
 DB 3661 TGCATTTGCTCGAGGCTGACATTTATCATATGAGACATGATCACTTACTCATGTTACAG 3720
 OY 3444 TGLTYRmetleugluGlyPheleuargserValCysleuGluasnGlyThrTPTThrSe 3464
 DB 3721 TGGATACATGTTGGAGGCTTCTCTGAGAGTGTGTTTGTAGAAAATGGACATGACATC 3780
 OY 3464 rProProlleCysAlaAlaValCysArgPheProCysGlnasnGlyGlyIleCysGlnAr 3484
 DB 3781 ACCCTCATTTTCAGAGCTGTCTGTGATTCATTCATGACAAATGGGGCATCTGCCAACG 3840
 OY 3484 GProAsnAlaCysSerCysProGluGlyTTPMetGlyArgLeuCysGluGluProlleCys 3504
 DB 3841 CCCAAATGCTTTTCTCTGTCGCA-GAGGGCTGGATGGGGCGCTCTGTGAGAACCAATCTG 3899
 OY 3504 sIleleuProCysleuasnGlyGlyValArgCysValAlaProTyrGlnCysArgProP 3524
 DB 3900 CATTCCTCCCTGCTGTAAGAGAGTGTGTGTGGCCCTTACCAAGTGTGACCTGCCGCC 3959
 OY 3524 OGlyTTPThrGlySerArgCysHisThrAla 3534
 DB 3960 TGGCTGGACGGGCTCTGCTGTCAATACAGCT 3990

RESULT 12
 AAK94919
 ID AAK94919 strand: cDNA; 4088 BP.

AAK94919:
 AC 06-NOV-2001 (first entry)
 DT Human full-length cDNA, SEQ ID NO: 4150.
 DE Human full-length cDNA; cDNA synthesis; oligo-capping; ss.
 KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
 XX Homo sapiens.
 OS
 XX EPI130094-A2.
 PN 05-SEP-2001.
 PD 07-JUL-2000; 2000BP-00114089.
 PF 08-JUL-1999; 993JP-00194486.
 PR 11-JAN-2000; 2000CP-00118774.
 PR 02-MAY-2000; 2000CP-00183765.
 XX (HELI-) HELIX RES INST.
 PA Ota T, Nishikawa T, Itoai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 DR WPI; 2001-524255/58.
 DR P-Psdb; AAM93953.
 XX 830 Primers useful for synthesizing full length cDNA clones and their use
 PT in genetic manipulation.
 XX Claim 8, SEQ ID NO 4150; 1380bp + Sequence listing; English.

CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been isolated
 CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesised by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the

CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 CC
 XX
 SQ Sequence 4088 BP; 1067 A; 939 C; 979 G; 1103 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 0 Length: 4088
 Score: 6950.00 Matches: 1232
 Percent Similarity: 99.20% Conservative: 1
 Best Local Similarity: 99.12% Mismatches: 10
 Query Match: 34.80% Indels: 1
 DB: 4 Gaps: 0

US-09-977-053-4 (1-3571) x AAK94919 (1-4088)

OY 2329 leuGluasnGlnleuValleuValleuThrThrGluValAlaThrPheSer 2348
 DB 1 ATGAAACCAAGCTAGTATTAAAGAGATTGACCAAGAGTAAAGATTGACATTTTC 60

OY 2349 CysleuGluGlyValleuGlnGlyProSerValleuValleuValleuValleuValleu 2368
 DB 61 TGTAAAGAAAGGAGATGCTGCAAGGCGCTGTCTGTCTGAAATGCTTGCATCCAGCAA 120

OY 2369 TTPAsnAspSerPheProValCysleuValleuValleuValleuValleuValleuValleu 2388
 DB 121 TGGATGACCTTCTTCTGTTGTAGATGTTCTTGTGACCCCACTCCCTTAATTTCC 180

OY 2389 PheGlyValProlleProSerSerAlaLeuHisPheGlySerThrValIleValIleValIle 2408
 DB 181 TTGGTGTCCCAATCTTCTTCTGCTCTTCAATTTGGAAGTACTGCAATATTTCTGT 240

OY 2409 ValGlyIlePhePheleuArgGlyAsnSerThrThrleuCysGlnProAspGlyThrTTP 2428
 DB 241 GTAGGTGGGTTTTTCTTAAGAGAAATTTTACCACTTGTGCAAGCTGCAAGTGGCCTGG 300

OY 2429 SerSerProleuProGluCysValProValGluCysProGlnProGluGluIleProAsn 2448
 DB 301 AGCTCTCCACCTGACAGATGTGTCAGTAAATGTCCCAACCTGAGAAATCCCAAT 360

OY 2449 GlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysleuPro 2468
 DB 361 GGAATCATTAATGTCAGAGCTTGGCTTAATCTGACAGACTCTTATCTGCAAGCA 420

OY 2469 GlyPheGluIleuValGlyAsnThrThrThrleuCysGlyGluAsnGlyIleSTPleuGly 2488
 DB 421 GGCCTTGAATTTGGGAAATATATACCACTTGTGGAATAATGCTGACTGCTTGA 480

OY 2489 GlyArgProThrCysAlaAlaIleGluCysleuValProGlySerGluIleleuAsnGlyIle 2508
 DB 481 GGAACCAACATGTAAAGCATGTAGTGTGTAACCAAGAGATTTGAAATGGCAAA 540

OY 2509 PheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPhe 2528
 DB 541 TTCTCTTACAGCACTTACATGACAGACAGACGCTTACTTGTCAACCAAGGCTTT 600

OY 2529 ArgleuGluGlyProSerAlaIleuThrCysleuGluThrGlyAspTyrAspValAspAla 2548
 DB 601 CGGCTGGAAGTCCCAAGTCTTGAACCTTTTGAAGACAGGTGATTTGGATGTGATGCC 660

OY 2549 ProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGlu 2568
 DB 661 CCAATCTGCAATCCATCCATGTGATTTCCCAACCAACCAATGAAATGGTTTGTGAA 720

OY 2569 GlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnVal 2588
 DB 721 GGTGAGATTACAGCTATGCTATGTCATATATCATATGATGCTTCCCTGGGTTTCAAGTG 780

OY 2589 AlaGlyHisAlaMetGlnThrCysGluGluSerGlyTTPSerSerIleProThrCys 2608
 DB 781 GCTGTGATATCCATGACAGCTGTGAAGTCAAGTGTCAAGTTCCATCCCAACCAATGT 840

OY 2609 MetProIleAspCysGlyleuProProHisIleAspPheGlyAspCysThrIleValleu 2628

841 ATGCCAATAGACGTGGCCCTCCCTCCATATGATTTGGAGACGTGCTAAACCAAA 900
2629 AAPPAGSINGLYTHRPHEGLUGINGIWAAPAMETMEGLUVALPROLYTHRPRO 2648
901 GATGACACAGGATATTTAGAGAGAGACATATGAMCTCATATGTGACTCT 960
2649 HISPROFROTHRISEUGLYALVALALYSTRIPGLUASANTHYPSGLUSEPRO 2668
961 CACCCCTCTTATCATTTGGAGAGAGCGGCTAAACCTGGGAAAAATCAAGAGGTCTCCT 1020
2668 ALATHRHSERSEASAPHELEUTYRGLYTHMERCVALSERLYTHRCSYASAPROGLY 2688
1021 GCTACACATTCATCAAACTTTGTATGATACCATGTTTCATACACCTGATATCCAGGA 1080
2688 TYRGULUENUGLYASAPROVALLEUIECYSGINGIUSAPGLYTHRTPASNGLYSER 2708
1081 TATGAACTTGGGGAACCTGTGATCTCCAGAGAGATGAACTTGGAAATGGCAAT 1140
2709 ALAPROSECYSEIIESEIIEGLUCYASAPLEUPROTHRALAPROGLUASNGLYPHELEU 2728
1141 GCACCATCTCGATTCATTTGATGATGATCTTCCCTACTGCTCTGAAAATGGCTTTTG 1200
2729 ARGPEPTHGLUTHSERMEGLYSERVALVALGLITYRSECYSLYSPROGLYHISGLIE 2748
1201 CGTTTACAGACATGACATGGAGACTGCTGCTCAGATATAGCTGTAACCTGACACAT 1260
2749 LEUALIAGLYSERASPLEUPROLEUCYSEUGLYASAPLYSTPSEGLYALASERPRO 2768
1261 CTAGCAGGCTGACCTTAAGCTTGTCTAGAGAAATGAAAGTGGTGTGCTCCCA 1320
2769 ARGCSGLUALAIIESECYSLYSLYSPROASAPROVALMECASNGLYSERILEYSGLY 2788
1321 CGCTGGAAAGCCATTTTCATGCAAAAAGCCAAATCCAGTCAATGATCCATCAAAAGGA 1380
2789 SERASNTYRTHRYLEUSERTHYLEUTYRGLUCYASAPPROGLYTYRVALLEUASNT 2808
1381 ACCAATCAACATACCTGAGACAGTGTACTATGAGTGTGCCCCGAGATATGTGCTGAAT 1440
2809 GLYTHRGUARGARTHRCSYGLINASAPPLYASNTTPASPGLUASPGLUPROILECY 2828
1441 GGCACCTGAGAGAGAACATGCCAGATGACAA-ACCTGGATAGAGATGAGCCATTTGC 1499
2829 ILEPROVALASP CYSESERSEPROVALISERVALASNGLYGLNALARGIYASPGLU 2848
1500 ATTCCTGTGACATGCACTTCACCCCCAGCTCAGCCAAATGGCCAGGTGAGAGAGACGAG 1559
2849 TYRTHRPHEGLUYSGLUIIEGLUTYRTHRCYASNGLYGLYPHELEUENUGLYAL 2868
1560 TACACATTCCAAAAAGAGATTGAAATCACTTCGAATGAAGGTTCTTGTGAGGAGGCC 1619
2869 ARGSEARVALCYSEUVALASNGLYSERTPSEGLYALATHRPROAPCYSEVALIPRO 2888
1620 AGAGAGTGGGTTGTCTTGCATGAGATGAGATGAGAGCACTCCCACTGTGAGCT 1679
2889 VALIARGCYSEALATHRPROFROGLUENUALASNGLYVALTHRGULUYLEUASPTYRGLY 2908
1680 GTCAAGATGGCAACCCGCCCAACATGGCCAAATGGGTGACGGAAGCCCTGAGCTATGGC 1739
2909 PHEMETLYSGIUALATHRPHENISCYSHISGLUGLYTYRILEUENISGLYALAPOLY 2928
1740 TTCATGAGAGAGATACATTTCCACTGTCAAGGGCTACATCTTTCAGCAGTGTCCAAA 1799
2929 LEUTHRCYSGINSEARPGLYASNTTPASPAJAGLUILEPROLEUCYSLYSPROVALAAN 2948
1800 CTCACCTGTGACATGACATGCAACTGGAGATGCAAGATTCCTCTCTGTAACCAAGCAAC 1859
2949 CYSGIYPROFROGLUASPLEUALAHISGLYPHEPROASNGLYPHESETHIEHISGLY 2968
1860 TGTGAGACCTCTGAGAGATCTTGCCCAATGGTTTCCATATGATTTCTTATATCAAGG 1919
2969 GLYHISILEGLINTYRGINCYSEHETROGLYTYRGLYLEUENISGLYASNSERSEARAGAG 2988

1920 GGCATATATACATATGATGCTTCTCGTTATAGCTCCATGGAATTCATCAAGAG 1979
2989 CYLEUSERASNGLYSERTRPSEGLYSESERPROSECYSELEUPROCYASARGYSE 3008
1980 TGCCTTCCAAATGGCTCTGGAGTGGCACTCATCTTCCGCTGCTTGCATATGTTCC 2039
3009 THPROVALILEGLUTYRGLYTHRVALASNGLYTHRASPHASPCYSGIYVALA 3028
2040 ACACAGATATTAATGAACTGTCAATGGAGCATTTTTCATCTGTGAAAGGCGCC 2099
3029 ARGILEGLUCYSEPHELYSGLYPHELYSEUGLYLEUSERGULIETHRCYSGIUAL 3048
2100 CGATTCAGTGTCTCAAGGCTTCAAGCTCTAGAGACTTCTGAATCACTGTGAAGCC 2159
3049 ASPGIYGLNTPSESERGLYPHEPROHISCYSGIULHISSTRHCYSGIYSELEUPRO 3068
2160 GATGGCCAGTGGAGCTCTGGTTCCCCCACTGGAACAACCTTCTGTGTCTCTTCCA 2219
3069 METILEPROASNALAHEIIESEGLIUTHRSESETRIPLYSGIUBANVALIETHRYR 3088
2220 ATGATACCAATGCGTTTCATCAGTGAACCACTCTTGAAAGAAATGATTAATCTTAC 2279
3089 SERCYARGSERGLYTYRVALIIEGLINGLYSESETRIPLYSGIULYSGLY 3108
2280 AGCTGAGGTCTGATATGATCAAGGCAAGTTCAGATTCAGTGTGACAGAAAGGG 2339
3109 VALTRPSEGLIAPROLYPROVALCYSGIUPROLEUSERCYSGIYSEPROSESERVAL 3128
2340 GTATGAGCCAGCTTATCCAGCTGTGACCTTCTGTGTGGTCCCAACGCTGTC 2399
3129 ALASNALAVALATHRGIYGLUALAHISHTYRGIUSERGULVALYSELEUARGCY 3148
2400 GCCAATGCACTGGCACTGGAGAGGACACACCTATGAAGTCAAGTCAACCTGATGT 2459
3149 LEUGIUGLYTYRTHMETASPTHRASPTHRASPTHRPHETHRCYSGILYASAPGLYARG 3168
2460 CTGAAAGTTATACATGATACATGATACATGATACATGATACATGATACATGATGTC 2519
3169 TRPHEPROGLUARGIIESECYSESPROLYSECYSEPROLEUPROGLUASNTIETHR 3188
2520 TGGTTCCCTGAGAGATCTCTGCACTCTTCAAAAATGCTCTCTCCGGAACATPACA 2579
3189 HIEILEUVALHISGLYASAPPHESERVALASNTGILVALSERVALSERCYSEALA 3208
2580 CATATCTTGTACATGAGACATTTCAAGTGAATGAAGCAATTTCTGTGATGTGCA 2639
3209 GLUGLYTYRTHRPHEGLIYVALAENIIESEVALCYSEGLUENUSAPGLYTHRTPGLU 3228
2640 GAAGGATATATCTTTAGAGATTAATCAATATGATATGATGATGATGATGATGATG 2699
3229 PROPROPHESERASPGIUSERCYSEPROVALISERCYSGIYLYAPROGLIUSERPROGLU 3248
2700 CCACCATTTCCGATGAATCTTGGACGTCCAGTGTCTTGTGGAAACCTGAATGCCAGAA 2759
3249 HISGLYPHEVALVALGLYSELYSTYRTHRPHEGLUSERTHRIELEIETHRYGICYSGIU 3268
2760 CATGATTTGTGTGTGGAGTAATACACCTTGAAGACAAATTTATATCAGTGTGAG 2819
3269 PROGLIYTYRGIUSLEUGLYASNTARGIUALARGVALCYSEGLIUSANARGINTPSE 3288
2820 CTGGCTATGAACTAGAGGAGAACAGGAACTGTCTCCAGAGAAACAGACAGTGAAT 2879
3289 GLIYGLYVALAIIECYSELYSGIUTHRARGYSGIUTHRPROLEUGIUPHELEUASNGLY 3308
2880 GGAGGGGTGCATATGCAAAAGAGACAGGTGGAATCTCACTTAATTTCCATATGGG 2939
3309 LYBALAASPTILEGLUASNTARGTHRGLYPROASNTVALIYTYRSECYSEASNTARGIY 3328
2940 AAAGCTGACATTTGAAACAGAGACACTGAGACCCAAAGTGTATATTCGCAACAGAGGC 2999
3329 TYRSELEUGIUGLYPROSEGLUALAHISCYSTHRGILUASNGLYTHRTRPSEHISPRO 3348
3000 TACAGCTTGAAGGCGCATCTGAGGACACTGACAGAAATGGAACCTGGAGCCACCA 3059

QY	2676	LeuTYrGlyYThrMeValSerTYrThrCysAsnProGlyYrGluLeuLeuGlyAsnPro	2653
Db	731	CTGATGGGTACCAATGGTTTCATACCCGTATATCCAGATATAGAACTTCTGGGAACCT	790
QY	2696	ValIleuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIle	2715
Db	791	GTCCTGACTCTGCCAGGAAGATGGAACTTGGAAATGGCAATGCCATCTTGCAATTTCATT	850
QY	2716	GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMet	2735
Db	851	GAATGTGACTTGCCACTGCTCTCCATAAATGGCTTTTGGCTTTTACAGAGACTTAGCAATG	910
QY	2736	GlySerAlaValAlaGlnTYrSerCysGlySerProGlyHisIleLeuAlaGlySerAspLeuArg	2755
Db	911	GGAAATGCTGTGCAGATATATAGCTGTAAACCTGCACATTCATAGCAGGCTCTGCATTAAGG	970
QY	2756	LeuCysLeuGluAsnArgGlyTrpSerGlyAlaSerProArgCysGluAlaIleSerCys	2775
Db	971	CTTTGTCTAGAGAAATAGAAAGTGGAGTGTGCTCTCCACGCTGTGAAGCCATTTCAATGC	1030
QY	2776	LysIleProAsnProValMetAsnGlySerIleLysGlySerAsnTYrThrTYrLeuSer	2795
Db	1031	AAAAAGCCAAATCCAGTCAATGAATGATCCATCAACAAAGAGAACTACATCACTATCTAGC	1090
QY	2796	ThrLeuTYrTYrGluCysAspProGlyTYrThrValLeuAsnGlyYthrGluArgThrCys	2815
Db	1091	ACGTTGTACTTGATGTGAGCCCGGATATGTCTGAAATGGCACTGAGAGAGAAACAATGC	1150
QY	2816	GluAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer	2835
Db	1151	CAGATGACAAATCTGGAGATGAGATAGCCCATTTTCATTTCTGTGACTGCAATTCA	1210
QY	2836	ProProValSerAlaAsnGlyGlnValArgGlyAspGluTYrThrPheGlnLysGluIle	2855
Db	1211	CCCCAGCTCAGGCCCAATGCGCCAGGTGAGAGAGACGATACATTCCAAAAAGAAATT	1270
QY	2856	GluTYrTYrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla	2875
Db	1271	GAATCACATTCACATGAAGGGTTCTTCTGAGGAGGCACAGATCGGGTTGTCTTGCC	1330
QY	2876	AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
Db	1331	AATGAACTGTGAGAGAGACCACTCCAGCTGTGTGCTGTCAAGTGTGCCACCCGCCA	1390
QY	2896	GlnLeuAlaAsnGlyValThrGluGlyLeuAspTYrGlyPheMetLysGluValThrPhe	2915
Db	1391	CACTGGCCAAATGGGTGACCGAAGGCTGTGACTATGGCTTCATGAAGAGATAACTTC	1450
QY	2916	HisCysHisGluGlyTYrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGly	2935
Db	1451	CACGTCAACGAGGGCTACATCATCTTGACCGGTCTCCAAAACCTCACCTGCACATCGAATGC	1510
QY	2936	AsnTrpAspAlaGluIleProLeuCysLysProValAsnGlyLysProProGluAspLeu	2955
Db	1511	AACGGAGAGCAAGAAATCTCTCTGTAAACACGCAACTGTGACCTCTGAAAGATCTT	1570
QY	2956	AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTYrGlnCys	2975
Db	1571	GCCCAATGTTTCCCTAATGGTTTTTCTTATTATCATGGGGCCATATACATACAGTGC	1630
QY	2976	PheProGlyTYrTYrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp	2995
Db	1631	TTTCTCGGTATTAAGCTCCATGTGAATTCATCAAGAAAGTGCCTCTCCAAAGGCTCTCTGG	1690
QY	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTYrGly	3015
Db	1691	AGTGGACGCTACCTCTGCTGCTGTGCAGATGTTCACACCAAGATATTAATATAGA	1750
QY	3016	ThrValAsnGlyYthrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly	3035
Db	1751	ACTGTCAATGGGACAGATTTTGACTGTGGAAAGCAAGCCCGAATTCAGTGTCTTCAAAGGC	1810
QY	3036	PheLysLeuLeuGlyLysSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly	3055

Db	1811	TTCAAGCTCTTAGAAGCTTTCTGAATACCTGTGAAGCCATGACCGATGAGCTCTGGG	1870
Oy	3056	PheProHisCysGluHisIleThrSerCysGlySerLeuProMetIleProAsnAlaPheIle	3075
Db	1871	TTCCCCACGTGTGAACAACATCTTGTGGTGTCTCTTCCAAATGATACCAAAATGCGTTGATC	1930
Oy	3076	SerGluThrSerSerTTPdlyGluAsnValIleIleThrTyrSerCysAlaGserGlyTyrVal	3095
Db	1931	AGTAGACACGAGCTTTGGAAAGGAAAAGTGAATACCTTACAGCTGCAGGTCTGATGATGTC	1990
Oy	3096	IleGlnGlySerSerAspLeuIleCysThrGluGlyGlyValTTPSerGlnProTyrPro	3115
Db	1991	ATACAGAGCACTTCAAGATCGAATTTGATTCAGAGAAAGGGATATGAGCCAGCTTATACCA	2050
Oy	3116	ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly	3135
Db	2051	GTCTGTGAGGCCCTTGTCTGTGGGGTCCCAACCGTCTGTCCGCAATGCAGTGGCAATCGGA	2110
Oy	3136	GluAlaHisIleThrTyrGluSerGluValIleLeuArgCysLeuGluGlyTyrThrMetAsp	3155
Db	2111	GAGGCACACACTTATGAAAGTGAAGTAAACTCAGATGCTCGAAGAGGTATACATGATGAT	2170
Oy	3156	ThrAspThrAspThrPheThrCysGlnIleAspGlyArgTTPheProGluIleSer	3175
Db	2171	ACAAATACAGATACATTCACCTGTCAGAAAGATGTCGCTGGTTCCTTGAGAGATCTCC	2230
Oy	3176	CysSerProIlyAslyCysProLeuProGluHisIleThrHisIleLeuValHisGlyAsp	3195
Db	2231	TGCAGTCTTAAAAATATCTCTCTCCGGGAAAAACATTAACATATACCTTGTACATGGGAC	2290
Oy	3196	AspPheSerValAsnArgGluValSerValSerCysAlaGluGlyTyrThrPheGluGly	3215
Db	2291	GATTTCAGTGTGAATAGGCAAGTTCTGTGTCTCATGTGCAGAGAGGTATACCTTGTAGAGGA	2350
Oy	3216	ValAsnIleSerValCysGlnLeuAspGlyThrThrGluProProPheSerAspGluSer	3235
Db	2351	GTTAACTATACAGATATGTACAGCTTGAAGAACCTGGAGACCAATCTCCGATGAATCT	2410
Oy	3236	CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer	3255
Db	2411	TGCAGTCAAGTTCTTGTGGGAAACCTGAAGTCAAAACATGATTTGTGTGGCAGT	2470
Oy	3256	LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGly	3275
Db	2471	AAATACACCTTTGAACACAAATATATTATACAGTGTAGGCTGGCTATGAACTAGAGGGG	2530
Oy	3276	AsnArgGluArgValCysGlnGluAsnArgGluThrSerGluGlyValAlaIleCysLys	3295
Db	2531	AACGGGAAACGTGTCTCCACGGAGAAACAGACAGTGAAGTGAAGGGGTGGCAATATGCAAA	2590
Oy	3296	GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg	3315
Db	2591	GAGACACAGGTGTGAACATCCACTTGAATTTCTCAATGGAAAGCTGACATTTGAAAAACAG	2650
Oy	3316	ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer	3335
Db	2651	AGGACTGGAACCCAAACGTGTATATTCTGTCAACAGAGGCTTACAGTCTTGAAGGGCCATCT	2710
Oy	3336	GluAlaHisIleCysThrGluAsnGlyThrTTPSerHisProValProLeuCysLysProAsn	3355
Db	2711	GAGGCACACATGCAAGAAATGGAACCTGGAAGCCACCACTCCCTCTCTCAAAACCAAT	2770
Oy	3356	ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr	3375
Db	2771	CCAGTCCCTGTCTCTTTGTGTGATTTCCGGAATGCTCTGTCTGTAAGAGAGATTAT	2830
Oy	3376	ValAspGluAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyTle	3395
Db	2831	GTTGATCAGATATGTCCATCAATGTAGGGAAGGTTTCTGTGTCAGGGCCACCGCATC	2890
Oy	3396	IleThrCysAsnProAspGluThrTTPThrGlnThrSerAlaLysCysGlyLysIleSer	3415

DB 2891 ATTACCTGCACCCCGACGACGCTGACACAGACCAAGCCCAATGTGAAAAATCTCA 2950
QY 3416 CysGlyProProAlaHisValGluAmaAlaIleAlaArgGlyValHisGlyGlnTyrGly 3435
DB 2951 TGTGGTCCAGCACTCAGTAGAAGAAATGCAATTGCTCGAGCGCTACATTATCAATTGGA 3010
QY 3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
DB 3011 GACATGATACCACTACATGTTACAGTGATACATGTTGAGAGGTTTCCGAGAGAGTGT 3070
QY 3456 CysLeuGluAmaGlyThrTyrThrSerProProIleCysArgAlaValCysArgPhePro 3475
DB 3071 TGTTTAGAAATGGAACATGACATCACTCTATTGTCAGACCTGCTGTGCAATTTCCA 3130
QY 3476 CysGluAmaGlyGlyIleCysGluArgProAlaCysSerCysProGluGlyTyrPheMet 3495
DB 3131 TGTCAAAATGAGGGGCACTGCGCAAGCCCAAAATGCTGTTCCTCCAGAGGCTGAGTANG 3190
QY 3496 GLyArgLeuCysGluGluProIleCysIleLeuProCysLeuAmaGlyGlyArgCysVal 3515
DB 3191 GGGCGCTCTGTGAAGAACCAATCTGCTTCTCCGTGTGAACGAGGCTGCTGTG 3250
QY 3516 AlaProTyrGlnCysArgProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
DB 3251 GCCCTTACAGTGTAAGTGCCTGCGCTGAGCGGAGTCTGCTGTGATACAGCTGTT 3310
QY 3536 CysGlnSerProCysLeuAmaGlyGlyIleCysValArgProAlaCysHisCysLeu 3555
DB 3311 TGCACATCTCCCTGCTAAATGTGGAATGTGAACCAACCATGTCACTGCTT 3370
QY 3556 SerSerTyrThrGlyHisAmaCysSerArgLyArgArgThrGlyPhe 3571
DB 3371 TCTTCTGACGGGACATTACTTCTCCAGAAAAGAGAGACTGGTTT 3418
RESULT 14
AAS26868
ID AAS26868 standard; cDNA; 3804 BP.
XX
AAS26868:
XX
07-NOV-2001 (first entry)
XX
Human cDNA encoding a novel secreted protein, SEQ ID 60.
XX
Human: immunosuppressive; antiarthritic; aa: antirheumatic; cytostatic;
XX cardiact; vasotropic; cerebroprotective; nootropic; neuroprotective;
XX antibacterial; virucide; fungicide; ophthalmological; vulnary;
XX secreted protein; rheumatoid arthritis; hyperproliferative disorder;
XX cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
XX cerebral ischemia; angiogenesis; nervous system disorder;
XX Alzheimer's disease; infection; ocular disorder; corneal infection;
XX wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
XX
Homo sapiens.
XX
MO200155441-A2.
XX
02-AUG-2001.
XX
17-JAN-2001; 2001MO-US001320.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180638P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0190874P.
PR 17-MAR-2000; 2000US-019076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218390P.
PR 26-JUL-2000; 2000US-0220563P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 06-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 14-SEP-2000; 2000US-0233066P.
PR 21-SEP-2000; 2000US-0234274P.
PR 21-SEP-2000; 2000US-0234275P.
PR 25-SEP-2000; 2000US-0234998P.
PR 25-SEP-2000; 2000US-0234999P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236378P.
PR 29-SEP-2000; 2000US-0236379P.
PR 29-SEP-2000; 2000US-0236380P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246177P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
PI WPI; 2001-476222/51.
DR P-PSDB; AAU16963.
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, for treating blood clotting disorder,
PT hemophilia.
XX
XX Claim 1; SEQ ID NO 60; 601bp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in

CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, anglogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The

Alignment Scores:
Pred. No.: 8e-315 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.91% Mismatches: 0
Query Match: 32.30% Indels: 0
DB: 4 Gaps: 0

US-09-977-053-4 (1-3571) x AAS26868 (1-3804)

QY 2436 ValProValGluCysProGluProGluGluLeuProGluGluLeuAspValGluGly 2455
DB 11 GTTCAGTGAATGATGCTCCCAACCTGAGAAATATCCCAATGAAATGATGATGCAAGGC 70

QY 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGluPheGluLeuValGluAsn 2475
DB 71 CTTCGCTATTCAGACAGAGCTCTCAATCCGCAAGCCAGGCTTGAATGGTGGAAAT 130

QY 2476 ThrThrThrLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 2495
DB 131 ACTACACACCTTTGTGAGAAATGTCATGCTGTGAGAGAAACCAACATGTAAGCC 190

QY 2496 IleGluCysLeuLysProLysGluLeuLeuAsnGluLysPheSerTyrThrAspLeuHis 2515
DB 191 ATTGAGTCCCTGAACCAACCAAGAGATTTGATGCAATTCCTTACACGACCTTACAC 250

QY 2516 TyrGluGluThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla 2535
DB 251 TATGAGACAGACGTTACTACTTCTTGCAACCGAGCTTCCGCTGGAAGTCCAGTCC 310

QY 2536 LeuThrCysLeuGluThrGluAspTyrAspValAspAlaProSerCysAsnAlaLeuHis 2555
DB 311 TTGACCTGTTTGAAGACAGGATGATGGATGATGATGCCCAATCTTGCAATGCCATCCAC 370

QY 2556 CysAspSerProGluProIleGluAsnGlyPheValGluGluAlaAspTyrSerTyrGly 2575
DB 371 TGTGATTCCTCCCAACCAACCAATGATTTTGTGAAAGGTGCAATTAACGATAGGT 430

QY 2576 AlaIleIleIleTyrSerCysPheProGlyPheGluValAlaGlyHisAlaMetGluThr 2595
DB 431 GCCAATATCATCTACATGATGCTCCCTGAGGTTTCAGTGGCTGCTCATGCCATGACACC 490

QY 2596 CysGluGluSerGlyTyrThrSerSerSerIleProThrCysMetProIleAspCysGluLeu 2615
DB 491 TGTAAAGGTGAGATGCTCAAGTTCATCCCAATGATTAAGCAATGATGATGATGATGATG 550

QY 2616 ProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGluGlyTyrPheGlu 2635
DB 551 CTTCTCATATGATGATTTTGAAGATGATTAACCAAGATGACCAAGATATTTGAG 610

QY 2636 GlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHisLeuGly 2655
DB 611 CAAGAAGACGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670

QY 2656 AlaValAlaLysThrTyrGluAsnThrLysGluSerProAlaThrHisSerSerAsnPhe 2675

Db	671	GCAGTGGCTAAAACTGGGAAATACAAAAGGAGTCTCCGTACACATTCACAAACTTT	1751	ACTGCAATGGACAGATTTTGACTGTGGAAAGCAGCCCGAATTCAGTCTCAAAAGCC	Db
Qy	2676	LeuTYRGIYThrMetValSerTYrThrCysAsnProGlyTYRGIYLeuLeuGlyValAsnPro	3036	PheLYSLeuLeuGlyLYLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly	Qy
Db	731	CTGATGGATCCATGATGTTTCTATACACCTGTATACCAAGATATGAACCTCTGGGAAACCTT	1811	TTCAAGCTCCAGAGACTTTCTGAAATACACCTGTATACCAAGGATGCGATGGAGCTCTGGG	Db
Qy	2696	ValIleuIleCYSGlnGlyValAspGlyTYrThrTrpAsnGlySerAlaProSerCysIleSerIle	3056	PheProHISCYSGlnHISThrSerCYSGlySerLeuProMetIleProAsnAlaPheIle	Qy
Db	791	GGCTGATCTGCAAGAAAGATGGAACTTGGAATGGAGATGGACAACTCCGACATTCCAATT	1871	TTCCCCCACTGTGAACACACTTCTGTGGTGTTCTCTCCAAATGATACCAAAATGCGTTCATC	Db
Qy	2716	GIYCYAsnProLeuProThrAlaProGlyValAsnGlyPheLeuValArgPheThrGluThrSerMet	3076	SerGIYThrSerSerTrpLYSValAsnValIleThrTYrSerCysAsnSerGlyTYrValI	Qy
Db	851	GAATGTGACTGCTCCATCTGCTCTCAAAAATGCGCTTTTGGCTTTTACAGAGACTGACATG	1931	AGTAGACACACTCTTGGAAGAAAAGTGAATTACTTACAGCTCAGAGTCTGATATATGTC	Db
Qy	2736	GIYSerAlaValAlaGlnTYrSerCysLYSProGlyHISIleLeuAlaGlySerAspLeuArg	3096	IleGlnGlySerSerAspLeuIleCYSThrGlyValTYrSerGlnProTYrPro	Qy
Db	911	GGAGTGGCTGTGCAGATATAGCTGTAAACCTGGACACATTTTACAGAGGCTGTGACTTAAAG	1991	ATACAGGACATTCAGATCTGATTTGTATGACAGAAAGGGATATGAGCCAGCTTATCCA	Db
Qy	2756	LeuCYLeuGluAsnArgLYeTrpSerGlyAlaSerProArgCysGluValAlaIleSerCys	3116	ValCYGlnProLeuSerCYSGlySerProProSerValAlaAsnAlaValAlaThrGly	Qy
Db	971	CTTGTCTATGAAATATGAAGATGGAGTGGTGGCTCCCAACGCTGTGAAGGCCATTTCAATGC	2051	GTCGTGAGCCCTGTGCTGTGGGGTCCCAACGCTGTGCGCAATGAGTGGCAACTGGA	Db
Qy	2776	LYSLYSProAsnProValMetAsnGlySerIleLYSGlySerAsnTYrThrTYrLeuSer	3136	GluAlaHISThrTYrTYrGluSerGluValLYSLeuArgCYLeuGlyGlyTYrTYrThrMetAsp	Qy
Db	1031	AAAAAGCCAAATCCAGTCAATGATCATCAAAAGAGACAACTACATCACTGACGC	2111	GAGGACACACTATGAAGAGTAAATCAAGATGTCTGAAAGGTTATATACATGAT	Db
Qy	2796	ThrLeuTYrTYrGlyCysAsnProGlyTYrValIleuAsnGlyTYrGluValArgThrCys	3156	ThrAspThrAspThrPheThrCysGlnLYSAspGlyArgTrpPheProGluValIleSer	Qy
Db	1091	ACGTTGTACTAAGATGTGACCCCGATATATGTGCGAATGTGCATGAAGAGGAACATGTC	2171	ACAGATACAGATACATTCACCTGTCAAGAAATGTGCTGGTTCCTTGAGAGAAATCTCC	Db
Qy	2816	GlnAspAspLYSAsnTrpAspGlyValAspGluProIleCysIleProValAspCysSerSer	3176	CysSerProLYSLYSCysProLeuProGlyValAsnIleThrHISIleLeuValHISGlyAsp	Qy
Db	1151	CAGATGACAAAAACTGGGATAGAGATGAGCCCAATTTGCAATTCCTGTGACGTGCAATTC	2231	TGCAGTCTTAAAAATATGCTCTCCCGGAAAAACATACATATATCTGTATACATGGGAC	Db
Qy	2836	ProProValSerAlaAsnGlyGlnValArgGlyAspGlyTYrThrPheGlnLYSLeuIle	3196	AspPheSerValAsnArgGlnValSerValSerCYAlaGlyGlyTYrThrPheGlnGly	Qy
Db	1211	CCCCAGTCTCAGCCAAATGGCCAGGTGAGAGAGACAGATACACATTCCAAAAAGAGATT	2291	GATTCAGTGTGAATAGGCAAGTTCGTGTCAATGTCAAGAGGTTATCCTTGAGAGGA	Db
Qy	2856	GluTYrTYrThrCysAsnGlyCysPheLeuLeuGlyValAlaArgSerArgValCYLeuAla	3326	ValAsnIleSerValCYSGlnLeuAspGlyTYrThrGlnProProPheSerAspGlySer	Qy
Db	1271	GAATACACTTGCAAATGAAGGGTTCCTTGCTTGAGGAGCCAGAGATCGGGTTCCTTCC	2351	GTTAACATATCAGATATGACGTTGAATGAACTCGGAGCCACATTCCTCGATGAATCT	Db
Qy	2876	AsnGlySerTrpSerGlyValAlaThrProAspCysValProValArgCysAlaThrProPro	3326	CysSerProValSerCYSGlyLYSProGlySerProGlnHISGlyPheValIleGlySer	Qy
Db	1331	AATGAAAGTTGGAGTGGAGCCACTCCCACTGTGCTGTCAAGTGTGCCACCCCGCCA	2411	TGCAGTCAAGTTCTTGTGGGAAACGTGAAGTCCAAACATGGAATTTGTGGTGGCAGT	Db
Qy	2896	GlnLeuAlaAsnGlyValThrGlnGlyLeuAspTYrGlyPheMetLYSGluValIleThrPhe	3356	LYSTyrThrPheGlnSerThrIleIleTYrGlnCYSGluProGlyTYrGluLeuGlnGly	Qy
Db	1391	CAACTGGCCAAATGGGGGTGACGGAAGCCCTGACATATGCTTCATAGAAAGTATCATTC	2471	AAATACACTTGTGAACACAAATTTATATCAGGTGAGGCTGTGATGAATGAGAGGG	Db
Qy	2916	HISCYHISGlnGlyTYrIleLeuHISGlyValaProLYSLeuThrCYSGlnSerAspGly	3276	AsnArgGlnArgValCYSGlnGlyAsnArgGlnTrpSerGlyGlyValAlaIleCYLYS	Qy
Db	1451	CACGTGTACAGAGGGCTACATCTTGCACCGTGTCTCAAAACTCACTGTCAGTCAATATGGC	2531	AAACGGAGACGTGTCTGCAAGAGAACAGACAGTGGAGTGGAGGGGTGGCAATATGCAA	Db
Qy	2936	AsnTrpAspAlaGluIleProLeuCysLYSProValAsnCysGlyProProGluAspLeu	3396	GluThrArgCYSGlnThrProLeuGluPheLeuAsnGlyLYSAlaAspIleGluAsnArg	Qy
Db	1511	AACGTGGATGACGAATTCCTCTCTGTAAACCACTGTGACCTTCCTCGAAGATCTT	2591	GAGACGAGGTGTAAATCTCACACTTGATTTCTCAATGGAAAGTGCATTTGAAGAAACAG	Db
Qy	2956	AlaHISGlyPheProAsnGlyPheSerPheIleHISGlyGlyHISIleGlnTYrGlnCys	3316	ThrThrGlyProAsnValValTYrSerCYAsnArgGlyTYrSerLeuGlnGlyProSer	Qy
Db	1571	GCCCATGGTTTCCATATGGTTTTCTTTATTCATGGGGGGCATAATACAGATACAGTCC	2651	ACGACTGACCAACGTGTATATTCCTGCAACAGAGGCTACAGTCTTGAAGGGCCACTC	Db
Qy	2976	PheProGlyTYrTYrLeuHISGlyAsnSerSerArgArgCYLeuSerAsnGlySerTrp	3336	GluAlaHISCYThrGluAsnGlyTYrThrTrpSerHISProValProLeuCYLYSProAsn	Qy
Db	1631	TTTCTGTGTATTAAGCTCATCGAATAATCATCAAAAGGTGCTCTCCAAATGACCTCCGG	2711	GAGGACACTGACAGAAATGAAGCTGAGACCACTCAATCTCTCTGCAAAACCAAT	Db
Qy	2996	SerGlySerSerProSerCYLeuProCysArgCysSerThrProValIleGluTYrGly	3376	ProCYAspProValProPheValIleProGluAsnAlaLeuLeuSerGlyGlyPheTYr	Qy
Db	1691	AGTGCAGCTCACTTCTGCTGCTTGCAGATTTTCCACACCGATATTTGAATATATGGA	2771	CCATGGCCCTGTCTCTTTTGATTCCTCGAAGATGCTGTGCTGTAAGAGGATTTAT	Db
Qy	3016	ThrValAsnGlyTYrThrAspPheAspCYSGlyLYSValAlaArgIleGlnCysPheLYSGLY	3376	ValAspGlnAsnValSerIleLYSCysArgGlnGlyPheLeuLeuGlnGlyHISGlyIle	Qy
			2831	GTTGATACAGATGTGTCAATCAATGTATGAGAAAGTTTCTGTGCAAGGGCCACCGGATC	Db

Db	1751	ACTGTCAAATGGCAACAGATTTTGACTGTGGAAAGCAAGCCGGAATTCAGTGTCTCAAAAGCC	1810
OY	3036	PheIysLeuLeuGlnGlyLeuSerGlnIleThrCysGlnAlaAspGlyGlnTrpSerSerGly	3055
Db	1811	TTCAAGCTCCCTAGAGACTTTCTGAATACCTGTGAAGCCATGGCCAGCTCTGGG	1870
OY	3056	PheProHisCysGlyHisIleThrSerCysGlySerLeuProMetIleProAsnAlaPheIle	3075
Db	1871	TTCCCCCACTGTGAACAACATCTTGTGGTGTTCTCTTCCAAATGAATACCAAAATGCGTTCATC	1930
OY	3076	SerGlnThrSerSerSerTrpIleGlnAlaAsnValIleThrTyrSerCysAsnSerGlyTyrVal	3095
Db	1931	AGTGAACACACCTCTTGAGAGAAAATGTATTACTTACAGCTCAGGTCTGATATATGTC	1990
OY	3096	IleGlnGlySerSerAspLeuIleCysThrGlnIlySerGlyValTrpSerGlnProTyrPro	3115
Db	1991	ATACAGGCACTTACAGATCTGATTTGTACAGAGAAAGGGGTATGAGACGACCTTATCCA	2050
OY	3116	ValCysGlnProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly	3135
Db	2051	GTCGTGAGCCCTGTGCTCCTGTGGGTGCCACCGCTGTGCGCAATGCAGTGGCAACTGGA	2110
OY	3136	GlnAlaHisIleThrTyrGlnSerGlnValIlyIleLeuArgCysLeuGlnGlnGlyTyrThrMetAsp	3155
Db	2111	GAGGACACACCTATGAAAGTAAAGTAAATCTCAGATGTCTGAAAGGTTATACATATGAT	2170
OY	3156	ThrAspThrAspThrPheThrCysGlnIlyAsnAspGlyArgTrpPheProGlnIlyIleSer	3175
Db	2171	ACAGATACAGATACATTCACCTGTGACAAATATGTGCTGTGTTCCCTGAGAGAAATCTCC	2230
OY	3176	CysSerProIlyIlyCysProLeuProGlnIlyAsnIleThrHisIleLeuValHisGlyAsp	3195
Db	2231	TGCGATCCTAAAATGATTCCTCCCGGAAAACATTAACATATATCTGTATACATGAGGAC	2290
OY	3196	AspPheSerValAlaAsnArgIleValSerValSerCysAlaGlnGlyTyrThrPheGlnGly	3215
Db	2291	GATTCACGTGTGAATAGGCAAGTTCGTGTGCATGTCACAAAGGTTATCCTTGAGAGGA	2350
OY	3216	ValAsnIleSerValCysGlnIleAsnAspGlyThrTrpGlnProProPheSerAspGlySer	3235
Db	2351	GTTACATATCATGATATGTCAGCTGATGAGAACTGGAGGCCACCATTCCTCGATGAATCT	2410
OY	3236	CysSerProValSerCysGlyIlyAspProGlnSerProGlnHisGlyPheValIleGlySer	3255
Db	2411	TGCGATCAGATTTCTTGTGGGAAACCTGAAAGTCCAGAACATGATTTGTGGTGGCAGT	2470
OY	3256	IlySerTyrThrPheGlnSerThrIleIleTyrGlnCysGlnProGlyTyrGlnIleGlnGly	3275
Db	2471	AAATACACTTGTGAAGACAAATATTAATCAAGTGAAGCTGTGCTATGAACTAGAGAGG	2530
OY	3276	AsnAspArgIlyAspValCysGlnGlnAsnAspGlnTrpSerGlyGlyValAlaIleCysIly	3295
Db	2531	AACAGGAACTGTGCTGCGAGAGAACACAGACAGTGGAGTGGAGGGGTGGCAATTGCAAA	2590
OY	3296	GlnThrArgCysGlnThrProLeuGlnPheLeuAsnGlyIlyAsaIleAspIleGlnAsnArg	3315
Db	2591	GAGACCAAGGTGTGAACCTCCACCTTGAAATTTCTCAATGAGGAAGCTGACATTGAAAACAG	2650
OY	3316	ThrThrGlyProAsnValValTyrSerCysAsnAspGlyTyrSerLeuGlnGlyIlyProSer	3335
Db	2651	ACGACTGACCCCAACGTGTATATTTCTGCAACAGAGCTACAGTCTTGAGAGGCCATCT	2710
OY	3336	GlnAlaHisCysThrThrIlyAsnGlnTyrThrTrpSerHisProValProLeuCysIlyProAsn	3355
Db	2711	GAGGACACTGCAACAGAAATGAACTGAGGCCACCAATCTCTCTGCAAAACCAAT	2770
OY	3356	ProCysProValProPheValIleProGlnAsnAlaLeuLeuSerGlnIlyGlnIlyPheTyr	3375
Db	2771	CCATGGCCCTGTGCTTTTGTGATTCGCCGAGAAATCTGTGCTGTAAGAAAGGATTTAT	2830
OY	3376	ValAspGlnAsnValSerIleIlyCysArgGlnGlyPheLeuLeuGlnGlnIlyHisGlyIle	3395
Db	2831	GTTATATCAGATATGTCTCATCTCAATATGATGAGAAAGTTCCTGTGCGAGGGCCACCGGATC	2890

```
QY 3396 ILeThrCysAsnProArgIuTrpThrGlnThrSerAlaYsCysGluYleSer 3415
DB 2891 ATTAGCTGCAACCCGACGAGCGTGGACACAGCAAGCCGCAATGGAATAATCTCA 2950
QY 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgIleValHisTrpGlnTrpGly 3445
DB 2951 TGTGGTTCACACGACTCACTGAGAAATGCAATGCTGAGCGCTGACTTATCAATATGGA 3010
QY 3436 AspMetIleThrYsSerCysYsSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
DB 3011 GACATATATACCTTACATCATTTACAGTGATCAATGTGGAGGGTTCTCGAGAGTGT 3070
QY 3456 CysLeuGluAsnGlyThrTrpThrSerProProIleCysAlaArgAlaValCysArgPhePro 3475
DB 3071 TGTTTAGAAATGGAACATGACATCACTCTCTATTTGACAGGCTGTCTGCATTTCCA 3130
QY 3476 CysGlnAsnGlyIleGlyLeuGlnAsnArgProAsnAlaCysSerCysProGluGlyTyrMet 3495
DB 3131 TGTCAAGATGGGGGCACTTCCCAAGCCCAATGCTGTGTCTGTCCAGAGGCTGAGATG 3190
QY 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyIleArgCysVal 3515
DB 3191 GGGCGCTCTGTGAGAACCAATCTGCATTTCTCCCTGTCTGAAAGGAGTGTGTGTG 3250
QY 3516 AlaProTyrGlnCysArgCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
DB 3251 GCCCTTACAGAGTGTACTGCCCGCTGGCTGAGCGGGGTCTCGCTGTCAACAGTGT 3310
QY 3536 CysGlnSerProCysLeuAsnGlyIleGlyCysValArgProAsnAlaArgCysHisCysLeu 3555
DB 3311 TGCAGTCTCCCTGCTTAAATGTGGAAATGTGTAAAGACCAACCATGTCACTGTCTT 3370
QY 3556 SerSerTrpThrGlyHisAsnCysSerArgIleArgTrpGlyPhe 3571
DB 3371 TCTTCTTGACGGGACATTACTGTTCAGAAAGAGGACTGGTTT 3418

RESULT 15
ABA06548
ID ABA06548 standard; cDNA, 3804 BP.
AC ABA06548;
XX
XX 10-JAN-2002 (first entry)
XX
XX Human cDNA SEQ ID NO: 214.
XX
XX Human; gene therapy; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX proliferative disorder; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200154474-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001, 2001WO-US001349.
XX
XX 31-JAN-2000, 2000US-0179065P.
XX 04-FEB-2000, 2000US-0180628P.
XX 24-FEB-2000, 2000US-0184664P.
XX 02-MAR-2000, 2000US-0186350P.
XX 16-MAR-2000, 2000US-0189874P.
XX 17-MAR-2000, 2000US-0190076P.
XX 18-APR-2000, 2000US-0198123P.
XX 19-MAY-2000, 2000US-0205515P.
XX 07-JUN-2000, 2000US-0209467P.
XX 28-JUN-2000, 2000US-0214886P.
XX 30-JUN-2000, 2000US-0215135P.
XX 07-JUL-2000, 2000US-0216647P.
XX 07-JUL-2000, 2000US-0216880P.
```

```
PR 11-JUL-2000, 2000US-0217487P.
PR 11-JUL-2000, 2000US-0217496P.
PR 14-JUL-2000, 2000US-0218290P.
PR 26-JUL-2000, 2000US-0220963P.
PR 26-JUL-2000, 2000US-0220964P.
PR 14-AUG-2000, 2000US-0224518P.
PR 14-AUG-2000, 2000US-0224519P.
PR 14-AUG-2000, 2000US-0225213P.
PR 14-AUG-2000, 2000US-0225214P.
PR 14-AUG-2000, 2000US-0225266P.
PR 14-AUG-2000, 2000US-0225267P.
PR 14-AUG-2000, 2000US-0225268P.
PR 14-AUG-2000, 2000US-0225270P.
PR 14-AUG-2000, 2000US-0225447P.
PR 14-AUG-2000, 2000US-0225757P.
PR 14-AUG-2000, 2000US-0225758P.
PR 14-AUG-2000, 2000US-0225759P.
PR 18-AUG-2000, 2000US-0226279P.
PR 22-AUG-2000, 2000US-0226681P.
PR 22-AUG-2000, 2000US-0226868P.
PR 22-AUG-2000, 2000US-0227182P.
PR 23-AUG-2000, 2000US-0227009P.
PR 30-AUG-2000, 2000US-0228924P.
PR 01-SEP-2000, 2000US-0229287P.
PR 01-SEP-2000, 2000US-0229343P.
PR 01-SEP-2000, 2000US-0229344P.
PR 01-SEP-2000, 2000US-0229345P.
PR 05-SEP-2000, 2000US-0229509P.
PR 06-SEP-2000, 2000US-0229513P.
PR 06-SEP-2000, 2000US-0230437P.
PR 06-SEP-2000, 2000US-0230438P.
PR 08-SEP-2000, 2000US-0231242P.
PR 08-SEP-2000, 2000US-0231243P.
PR 08-SEP-2000, 2000US-0231244P.
PR 08-SEP-2000, 2000US-0231413P.
PR 08-SEP-2000, 2000US-0231414P.
PR 08-SEP-2000, 2000US-0232080P.
PR 08-SEP-2000, 2000US-0232081P.
PR 12-SEP-2000, 2000US-0231968P.
PR 14-SEP-2000, 2000US-0232397P.
PR 14-SEP-2000, 2000US-0232398P.
PR 14-SEP-2000, 2000US-0232399P.
PR 14-SEP-2000, 2000US-0232400P.
PR 14-SEP-2000, 2000US-0232401P.
PR 14-SEP-2000, 2000US-0233063P.
PR 14-SEP-2000, 2000US-0233064P.
PR 14-SEP-2000, 2000US-0233065P.
PR 21-SEP-2000, 2000US-0234274P.
PR 21-SEP-2000, 2000US-0234275P.
PR 25-SEP-2000, 2000US-0234997P.
PR 25-SEP-2000, 2000US-0234998P.
PR 26-SEP-2000, 2000US-0235484P.
PR 27-SEP-2000, 2000US-0235834P.
PR 27-SEP-2000, 2000US-0235836P.
PR 29-SEP-2000, 2000US-0236327P.
PR 29-SEP-2000, 2000US-0236328P.
PR 29-SEP-2000, 2000US-0236367P.
PR 29-SEP-2000, 2000US-0236368P.
PR 29-SEP-2000, 2000US-0236369P.
PR 29-SEP-2000, 2000US-0236370P.
PR 02-OCT-2000, 2000US-0237037P.
PR 02-OCT-2000, 2000US-0237038P.
PR 02-OCT-2000, 2000US-0237039P.
PR 02-OCT-2000, 2000US-0237040P.
PR 13-OCT-2000, 2000US-0239935P.
PR 13-OCT-2000, 2000US-0239937P.
PR 20-OCT-2000, 2000US-0240960P.
PR 20-OCT-2000, 2000US-0241221P.
PR 20-OCT-2000, 2000US-0241785P.
PR 20-OCT-2000, 2000US-0241786P.
PR 20-OCT-2000, 2000US-0241787P.
PR 20-OCT-2000, 2000US-0241808P.
PR 20-OCT-2000, 2000US-0241809P.
```

PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249247P.
 PR 17-NOV-2000; 2000US-0249248P.
 PR 17-NOV-2000; 2000US-0249249P.
 PR 17-NOV-2000; 2000US-0249250P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251038P.
 PR 05-DEC-2000; 2000US-0251039P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251889P.
 PR 08-DEC-2000; 2000US-0251907P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-476161/51.
 XX P-PsDB; ABB010326.
 XX
 PT Isolated nucleic acid molecule encoding an inflammation-associated
 PT polypeptide is used in preventing, treating or ameliorating a medical
 condition.
 XX
 XX Claim 1; SEQ ID NO 214; 859pp + Sequence Listing; English.
 CC The present invention provides human cDNAs, proteins and related genomic
 CC DNAs. These can be used in the treatment of neural, immune system,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal and proliferative disorders and inflammation. The present sequence
 CC is a cDNA of the invention
 XX
 SO Sequence 3804 BP; 1025 A; 858 C; 915 G; 1006 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	86-315	Length:	3804
Score:	6452.00	Matches:	1135
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.91%	Mismatches:	0
Query Match:	32.308	Indels:	0
DB:	4	Gaps:	0
US-09-977-053-4 (1-3571) x ABA06548 (1-3804)			
QY 2436 ValProValGluCysProGlnProGluGluValIleProAsnGlyIleIleLeuValGlnGly	2435		
DB 11 GTTCAAGTAAATGCTCCCAACCTGAAGAAATCCCAATGAAATCATTAATGTCAGAGCC	70		
QY 2456 LeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsn	2475		
DB 71 CTTCCTATCTCAGACACACTCTATATCTGAAACCGAGCTTGAATGGTGGAAAT	130		
QY 2476 ThrThrThrLeuCysGlyGluAsnGlyIleSTPLeuGlyGlyLysProThrCysLysAla	2495		
DB 131 ACTACACACCTTGTGAGAAATGCTCACTGCTTGAGGAAACCAACATGTAAAGCC	190		
QY 2496 IleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHis	2515		
DB 191 ATTGAAGTCCCTGAACCAACCAAGAGATTGAAATGCAAAATCTCTACACGAACTACAC	250		
QY 2516 TyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAla	2535		
DB 251 TATGACACAGACCTTATCTACTCTTCAACCGAGGCTTTCGCTCGAAGGTCCAGTGC	310		
QY 2536 LeuThrCysLeuGluGluThrGlyAspThrAspValAspAlaProSerCysAsnAlaIleHis	2555		
DB 311 TTGACCTGTTTGAAGACAGGTGATGGATGTGAAGTCCCACTTGCATATGCAATGCCATCCAC	370		
QY 2556 CysAspSerProGlnProIleGluAsnGlyPheValGluGlyValAspTyrSerTyrGly	2575		
DB 371 TGTGATTCCTCCCAACCAACCAATGAAGATTTGTGAAGAGTCCAGATTACAGTATAGT	430		
QY 2576 AlaIleIleIleTyrSerCysPheProGlyPheGluValAlaGlyIleSalameGlnThr	2595		
DB 431 GCCATATATCATCTACAGTGTCTCCCGGCTTCAAGTGGCTGTGATGCATGCAAGACC	490		
QY 2596 CysGluGluSerGlyTyrPheSerSerIleProThrCysMetProIleAspCysGlyLeu	2615		
DB 491 TGTGAAGAGTCAGAGATGTCACATTCATCCCAACATGATGCAATAGACTGTGGCTC	550		
QY 2616 ProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGlu	2635		
DB 551 CTTCTCATATATGATTTTGAAGCTGACTTAACTCAAAATGACAGGATATTTTGG	610		
QY 2636 GlnGluAspAspMetMetGluValProTyrValThrProHisProProTyrHisLeuGly	2655		
DB 611 CAAGAGACAGCATGATGAGAAATTCATATGATGATCTCAACCTCTTATCATTTGGGA	670		
QY 2656 AlaValAlaLysThrTyrGluAsnThrLysGluSerProAlaThrHisSerSerAspPhe	2675		
DB 671 GCAGTGTCTAAACCTGCGAAATAATACAAAGAGTCTCTCTCACTCATCACTT	730		
QY 2676 LeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnPro	2695		
DB 731 CTGTATGATGACATGATTTTGAAGCTGACTTAACTCAAAATGACAGGATATTTTGG	790		
QY 2696 ValLeuIleCysGlnGluAspGlyTyrThrAsnGlySerAlaProSerCysIleSerIle	2715		
DB 791 GTCTGATCTGCCAGAAAGATGAACTTGAATGCGAGTCAACATCTGATTTCAATT	850		
QY 2716 GluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThrGlnThrSerMet	2735		
DB 851 GAATGACACTGCTACTGCTCCGAAATGAGCTTTTGTGATTAAGAGACTAGGATG	910		
QY 2736 GlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArg	2755		
DB 911 GGAAGTCTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATG	970		

Dd	1031	AAAAAGCCAAATCCAGTCATGAAATGCAATCCATCAAAAGAGCACTACACATACCTGAGC	10930
Qy	2796	ThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyTyrGluIuArgTyrThrCys	2815
Dd	1091	ACGTTGTACTAATGAGTGTGACCCCGAGATATATGTGCTGAATGGCACTGAAGAGGAACATGC	1150
Qy	2816	GlnAspAspIlyAsnTTPAspGlnAspGluProIleCysIleProValAspCysSerSer	2835
Dd	1151	CAGGATGACAAAACCTGGGATGAGATAGACCATTTCATCTCGTGGACATGCAGTTCA	1210
Qy	2836	ProProValSerAlaAsnGlyGlnValAlaArgGlyAspAspIlyTyrThrPheGlnIleGluIle	2855
Dd	1211	CCCCAGCTCTCAGCCCAATGGCCAGGAGGAGAGAGACGAGTACATCTCCAAAAGAGATT	1270
Qy	1271	GAAATACACTTGCAATGAAGAGGTTCTTGCTTGAGGGAGCAGAGAGTCGGGTTGTCTGGC	1330
Qy	2876	AsnGlySerTTPSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
Dd	1331	AATGGAAGTTGGAGTGGAGCCACTCCCACTGTGCTGTGCAATGTGGCACCCCGGCA	1390
Qy	2896	GlnLeuAlaAsnGlyValThrGlnGlyLeuAspTyrGlyPheMetLysGlnValThrPhe	2915
Dd	1391	CAACTGGCCAAATGGGGGTGACGAGAGGCTGGACATATGGCTTCATGAGAGATTAATTTC	1450
Qy	2916	HisCysHisGlnGlyTyrIleLeuHisGlnAlaProLysLeuThrCysGlnSerAspGly	2935
Dd	1451	CACCTGTACAGAGGGCTTACATCTTGACAGGTGCTCCAAAATCACTGTGCATTCAGATGGC	1510
Qy	2936	AsnTTPAspAlaGlnIleProIleuCysLysProValAsnCysGlyProProGlnAspLeu	2955
Dd	1511	AACTGGATGACAGAAATTCCTCTCGTAAACACAGTCAACTGTGGACCTCCGAAAGATCTT	1570
Qy	2956	AlaHisGlnPheProAsnGlyPheSerPheIleHisGlnGlyHisIleGlnTyrGlnCys	2975
Dd	1571	GCCCAATGGTTCCCTTAATGGTTTTTCCCTTAATCATGGAGGACAAATACATATCAGAGC	1630
Qy	2976	PheProGlyTyrTyrLysLeuHisGlyValAsnSerSerArgArgCysLeuSerAsnGlySerTyr	2995
Dd	1631	TTTCTCGTTATAGCTTCATGGAAATTCATCAGAAAGGTGCTCTCCAAATGGCTCCGG	1690
Qy	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGlnTyrGly	3015
Dd	1691	AGTGGCACTCACTTCCCTGCTGCTTGACAGATGTTCCACACACAGTATGGAATATGGA	1750
Qy	3016	ThrValAsnGlyTyrThrAspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGly	3035
Dd	1751	ACTGTCAATGGGACAGATTTTGACTGTGGAAAGGAGCCGAGATTCAGTGTCTCAAAAGGC	1810
Qy	3036	PheLysLeuLeuGlyLeuSerGlnIleThrCysGlnAlaAspGlyGlnTTPSerSerGly	3055
Dd	1811	TTCCAGCTCTTCAGACTTCTTGAAATCACTGTGAAGCCGATGGCCAGTGAAGCTCTGG	1870
Qy	3056	PheProHisCysGlnHisIleThrSerCysGlySerLeuProMetIleProAsnAlaPheIle	3075
Dd	1871	TTCCCCCACTGGAAACACATCTTGTTGGTCTCTTCCAAATGATGCCAAATGGCGTTATC	1930
Qy	3076	SerGlnThrSerSerTTPLysGlnAsnValIleThrTyrSerCysArgSerGlyTyrVal	3095
Dd	1931	AGTGAAGACAGCTCTTGGAAGGAATATGATTAACCTTACACTGTGAGGTCTGGAATATTC	1990
Qy	3096	IleGlnGlySerSerAspLeuIleCysThrGlnLysGlyValATTPSerGlnProTyrPro	3115
Dd	1991	ATACAGAGCACTTCAGATCTGATTTGTACAGAGAAAGGGGTATGAGCCACGCTTATCA	2050
Qy	3116	ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly	3135
Dd	2051	GTCGTGAGCCCTTGTCTCTGTGGGTCCCCACGCTGTGTGCAATGAGTGGCAACTGGA	2110
Qy	3136	GlnAlaHisIleThrTyrGlnSerGlnValLysLeuArgCysLeuGlnGlyTyrThrMetAsp	3155
Dd	2111	GAGGACACACCTATGAAGAAGTGAAGTGAACCTCAATCTCTGGAAGGTTATACAGATGAT	2170

QY	3156	ThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluLysGlnIleSer	3175
Db	2171	AcGAGTAAcAGATAcATcATcCTcGcAcGAAATGcGTcGTcGTcCTcCGAcGAAATcTCC	2230
QY	3176	CysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp	3195
Db	2231	TGCAGTCTTAAAAATGTCCTCTCCCGSAAAAcATTAACAATATcCTGTATcATcGGGcAC	2290
QY	3196	AspPheSerValAsnAspGlnValIleSerValSerCysAlaGluGlyPyrThrPheGluGly	3215
Db	2291	GATTTCAcGTGcGAAATAGcCAcGATTTCTGTGTcATcAGTGTcAGAAAGGTATcATcCTTTGAAGGcA	2350
QY	3216	ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer	3235
Db	2351	GTTAAcATATcATcAGTATcGTcCAcCTGTATcGAAcCTcGGAcGcCAcCATTTCTCGAGATcCT	2410
QY	3236	CysSerProValIleSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer	3255
Db	2411	TGCAGTCAcAGTTTCTGTGTGGAAAcCTGAAGTcCAcGAAcATcGATATTTGTGTGGTGGcAGT	2470
QY	3256	LysPyrThrPheGluSerThrIleIleGlyGlnCysGluProGlyPyrTrpGluGluGly	3275
Db	2471	AAATAcACTTTGAAAGcCAATATTTATcTcAGTGTcAGCTGTcCTATcGAAcCTAcGAGGcG	2530
QY	3276	AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysLys	3295
Db	2531	AAcAGGAAcCTGTcCTcCGcAGAAcCAcGAcATcGAGTGTcGAGGcGTGTcGCAATTcGCAA	2590
QY	3296	GluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg	3315
Db	2591	GAGAcCAcAGTGTGAATcCTcCAcTGTGAATTTCTcCATcTGGAAAcGTcGAcATTGAAAcCAcG	2650
QY	3316	ThrThrGlyProAsnValValIleSerCysAsnArgGlyTrpSerLeuGluGlyProSer	3335
Db	2651	AcGAcTGAcCCcCAcAGcGTATcATTcCTcGcCAcCAcAGcCTAcATcCTTTGAAGcGCATcCT	2710
QY	3336	GluAlaHisCysThrGluAsnGlyPThrTrpSerHisProValProLeuCysLysProAsn	3355
Db	2711	GAGcCAcACTcCAcAGAAATcAGAAcCTcGAcGcCAcCAcAGTcCTcCTcCGcAAAcCAAAcT	2770
QY	3356	ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheThr	3375
Db	2771	CCATcCCCTGTcCTTTGTGTGATcCTcCGAcGATcCTcGTcGTcTcGAAAcGAcATTAT	2830
QY	3376	ValAspGluAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle	3395
Db	2831	GTTATcATcGATATcGTcCATcCAATcGTATcGGAAGcATcTTTCTcGTcGcAGGcGcCAcGGcATc	2890
QY	3396	IleThrCysAsnProAspGluThrTrpThrGlnHisSerAlaLysCysGluLysIleSer	3415
Db	2891	ATTAcCTGTcCAcCCcCGAcGAAcGTGTcGAcCAcGAcCAAcGcCAATcATcGAAAAATcCTcA	2950
QY	3416	CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisPyrGlnTrpGly	3435
Db	2951	TGTGTcCAcCAcAGTcCAcAGTcGAAATcGcATcCTcCTcGAcGGcCTcATcATcATcATcGGA	3010
QY	3436	AspMetIleThrTrpSerCysTrpSerGlyPyrMetLeuGluGlyPheLeuArgSerVal	3455
Db	3011	GAcATcGATcCAcCTTcACTcATcGTTTAcAGTGTcATcATcTGTcGAcGGcGTTCcTGTcGAcAGcGT	3070
QY	3456	CysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro	3475
Db	3071	TGTTTAGAAATcGAAcCATcGAcATcCAcCTcCTcATTcTGTcAGAcGTGTcCTGTcCATcTTcCA	3130
QY	3476	CysGlnAsnGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyPyrMet	3495
Db	3131	TGTcCAATcGAcGGGcCATcCTcGcCAAcGcCAAAcTcCTGTcCTcGTcCAcGAcGGcGTcGATc	3190
QY	3496	GluArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyLysArgCysVal	3515
Db	3191	GcGGcGcCTcCTGTcGAAAGcCAATcGTcATcTCTcCTcGTcTcGAAcGGAAGcGTcGTGTGTGT	3250

PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0255678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX MPI; 2001-581633/65.
DR P-PSDB; AAU87355.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
XX
PS Claim 1; SEQ ID NO 275; 837bp; English.
XX
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (III) encoded
XX by (I), are used to treat a medical conditions and in diagnosis of a
XX pathological condition. Disorders which are diagnosed or treated include
XX autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease and
XX amyotrophic lateral sclerosis, infections caused by bacteria, viruses
XX e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
XX e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
XX adenocarcinomas and irritable bowel syndrome, reproductive system
XX disorders e.g. testicular feminization, endocrine disorders e.g. diabetes
XX and pituitary dwarfism, cancers and disorders at the cellular level e.g.
XX leukemia, disorders involving neovascularization e.g. malignancies,
XX respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
XX acute kidney failure and blood related disorders e.g. myocardial
XX infarction. The polypeptides can also be used to aid wound healing and
XX epithelial cell proliferation, to prevent skin aging due to sunburn, to
XX maintain organs before transplantation, for supporting cell culture of
XX primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities, fat content, lipid, protein,

Alignment Scores:
Pred. No.: 8e-315 Length: 3804
Score: 6452.00 Matches: 1135
Percent Similarity: 100.00% Conservative: 1

Best Local Similarity:	99.91%	Mismatches:	0
Query Match:	32.30%	Indels:	0
DB:	4	Gaps:	0
US-09-977-053-4 (1-3571) x ABR43685 (1-3804)			
QY	2436	ValProValGluCyProGlnProGluGluIleProAenGlyIleIleAspValGlnGly	2455
DB	11	GTTCAGTGAATGTCCTCCCACTGAGAAATCCCAATGCAATCTTGAATGCAAGGC	70
QY	2456	IeuAlaIleThrLeuSerThrAlaLeuIleThrCysLeuProGlyPheGluIleuValGlyAen	2475
DB	71	CTTGCTATCTCAGACAGACTCTATACCTGCAAGCCAGGCTTGAATTTGGGGAAAT	130
QY	2476	ThrThrThrLeuCyGlyIleuAenGlyHisIlePLeuGlyGlyLeuProThrCysValAla	2495
DB	131	ACTACCACTCTTGTGGAGAAATGTCTGCTGCTGGAGAAACCAATGTAAAGCC	190
QY	2496	IleGluCysLeuLysProLysGluIleuAenGlyLysPheSerThrAspLeuHis	2515
DB	191	ATTGAGTGCCTGAACCCCAAGAGATTGTGAATGGCAATTCCTTACACGGACTACAC	250
QY	2516	ThrGlyGlnThrValThrIleSerCysAenAenGlyPheAenGluGlnIleProSerAla	2535
DB	251	TATGACAGACCTGTAACCTACTCTGCAACCAAGCTTGTGCTCAAGGCTCCAGTGC	310
QY	2536	LeuThrCysLeuGlyIleThrGlyAspIleAspValAlaPheSerCysAenAlaIleHis	2555
DB	311	TTGACCTGTTTGAAGACAGATGATGGATGGATGCCCATCTTGCAATGCCATCCAC	370
QY	2556	CysAspSerProGlnProIleGluAenGlyPheValGluGlyAlaAspIleSerThrGly	2575
DB	371	TGTGATTCCTCCCAACCCCACTGAAATGGTTTGTGAAGAGTCAAGATTAAGCTATGGT	430
QY	2576	AlaIleIleIleIleIleSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThr	2595
DB	431	GCCATATCATCTACAGATGCTCTCCCTGAGTTCAGTGGTATGCAAGCAGAC	490
QY	2596	CysGluGluSerGlyIleProSerSerIleProThrCysMetProIleAspCysGlyLeu	2615
DB	491	TGTGAAGACTCAGATGTCAGATTCATCCCAATGATATGCAATGATGCTGCTC	550
QY	2616	ProProHisIleAspPheGlyAspCysThrIleuLysAspAspGlnIleIlePheGlu	2635
DB	551	CTCTCTCATATGATTTTGGAGACTGTACTAACTCAAGATGACAGGATTTTGGAG	610
QY	2636	GlnGluAspAspMetMetGluValProIleValThrProHisProProIleGlnGly	2655
DB	611	CAAGAAGACGACATGATGGAATGATGATGATGATGATGATGATGATGATGATGATG	670
QY	2656	AlaValAlaIleThrThrIleuAenThrIleuGluSerProAlaThrHisSerSerAspPhe	2675
DB	671	GCGTGTGCTTAACCTGGAAATATCAAGAGTCTCCGTCAACATTCATCAAACTTT	730
QY	2676	IeuIleGlyIleThrMetValSerIleThrCysAenProGlyIleGluLeuLeuGlyAenPro	2695
DB	731	CTGTATGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	790
QY	2696	ValIleuIleCysGlnGluAspGlyIleThrTPAenGlySerAlaProSerCysIleSerIle	2715
DB	791	GGTGTGATGCTGCAAGAGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG	850
QY	2716	GluCysAspLeuProThrAlaProGluAenGlyPheLeuAenPheThrGluIleThrMet	2735
DB	851	GATGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	910
QY	2736	GlySerAlaValGlnIleSerCysLysProGlyHisIleLeuAlaGlySerAspLeuAenG	2755
DB	911	GGAGTGTGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	970
QY	2756	LeuCysLeuGluAspAspGlyIleThrSerGlyIleAspProAspCysGluAlaIleSerCys	2775
DB	971	CTTTGTCTGAAGATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA	1030

QY 2776 lyeLysProAsnProValMetAsnGlySerIleIysGlySerAsnTyrThrTyrLeuSer 2795
 DB 1031 AAAAAACCAATCCAGTCATGAATGATCCATCAAGAAAGCACTACACATACCTAGC 1090
 QY 2796 ThrLeuTyrThrGluCysAspProGlyTyrValIleuAsnGlyThrGluArgGlyThrCys 2815
 DB 1091 ACCTTACTATAGTGAAGTGAACCCCGATATGCTGATGAGCACTGAGAGAGAACTGC 1150
 QY 2816 GluAspAspLysAsnTrpAspGluAspGluProIleCysIleProValAspCysSerSer 2835
 DB 1151 CAGAGTACAAATAACGGATGAGATGAGATGCCATTCCTCTGATGATGAGTTCA 1210
 QY 2836 ProProValSerAlaAsnGlyIleValArgGlyAspGlyTyrThrPheGluIle 2855
 DB 1211 CCCCAGTCTCAGCCCAATGCGCAGTGAAGAGAGACAGTACATCTCAAAAAGAGATT 1270
 QY 2856 GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyValIleArgSerArgValCysLeuAla 2875
 DB 1271 GAATACACTTGCAATGAAGGTTCTTGCTGAGGAGGCCAGAGTGGGTTGTCTTGGC 1330
 QY 2876 AsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro 2895
 DB 1331 AATGGAAGTTGAGATGAGGACCATCCGACCTGTCTGTCAAGTGTGCCACCCGCCA 1390
 QY 2896 GluLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetGlyValThrPhe 2915
 DB 1391 CAACTGGCCCATGGGGTGAAGGAGGCTGAGACTATGGCTTCATGAAGAAAGTAACTTC 1450
 QY 2916 HisCysHisGluGluTyrTrileuHisGlyValaProIleuThrCysGlnSerAspGly 2935
 DB 1451 CACTGACAGAGGCTACATCTTGACGAGGCTCCAAACCTCCTCAGTCAAGTGAAGGC 1510
 QY 2936 AsnTrpAspAlaGluIleProIleProIleProValAsnCysGlyProProGluAspLeu 2955
 DB 1511 AACTGGAGTGCAGAGATCTCTCTGTAAACAGTCACTGTGAGACTCTCTGAAGATCTT 1570
 QY 2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGluTyrGlnCys 2975
 DB 1571 GCCCAGTGTTCCTTAATGTTTTCCTTATTCATGGGGGCCATATACAGTATCACTGC 1630
 QY 2976 PheProGluTyrTyrLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrp 2995
 DB 1631 TTTCCTGGTTAATAGCTCCATGGAATTCATCAAGAAAGTGCCTCTCAATAGGCTCTGG 1690
 QY 2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
 DB 1691 AGTGGAGCTCACTCTCTGCTGCTTGAGATTTCCACACCAAGTAATGAAATGGA 1750
 QY 3016 ThrValAsnGlyThrAspPheAspCysGlyValAlaAlaArgIleGluCysPheIleGly 3035
 DB 1751 ACTGTCATATGGGACATTTTGACTGTGAAGAGCCGCGATTCAGTCTTCAAAAGGC 1810
 QY 3036 PheIleValLeuGlyIleuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGly 3055
 DB 1811 TTCACCTCTTGAAGCTTCTGAAATCACTGTGAAGCCGATGAGCTCTGGG 1870
 QY 3056 PheProHisCysGlyHisIleThrSerCysGlySerLeuProMetIleProAlaAlaPheIle 3075
 DB 1871 TTCCCCCAGTGAACACACTTCTTGTTGTTCTCTTCATGATACCAATGGGTTTCATC 1930
 QY 3076 SerGluThrSerSerTrpIlyGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
 DB 1931 AGTGAACACAGCTCTTGAAGAAATGTGATTAATTAAGCTGACAGTCTGGAATAGTC 1990
 QY 3096 IleGluGlySerSerAspLeuIleCysThrGluIleGlyValTrpSerGlnProTyrPro 3115
 DB 1991 ATTCAGAGGAGTTCAGATCTGATTTGTACAGAAAGGATATGAGCCAGCTTATCCA 2050
 QY 3116 ValCysGluProIleuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly 3135
 DB 2051 GTCTGTGAAGCCCTTGTGTGTGGGTCCCAACGCTGTGTCCCAATGACAGTGAAGCACTGA 2110

QY 3136 GluAlaHisThrTyrGluSerGluValIleuLeuArgCysLeuGluGlyTyrThrPheAsp 3155
 DB 2111 GAGGCACACACTATGAAGAGTGAAGTGAACCTCAGATCTGGAAGGTTATACATGAT 2170
 QY 3156 ThrAspThrAspThrPheThrCysGluIleuAspGlyArgTrpPheProGluIleSer 3175
 DB 2171 ACAGATACAGATACATTCCTCTGTCAAGAAAGTGTGCTGTGTTCTCTGAGAAATCTCC 2230
 QY 3176 CysSerProIlyIysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
 DB 2231 TGCAGCTCTAAATAATGTCTCTCCGGAATAACATATATCTGTACATGGGCAC 2290
 QY 3196 AspPheSerValAsnArgGluValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
 DB 2291 GATTTCAGTGAATAGGCAAGTTCGTGTGCATGTGCAGAAAGGATATACCTTGAAGGGA 2350
 QY 3216 ValAsnIleSerValCysGluIleuAspGlyThrTrpGluProProPheSerAspGlySer 3235
 DB 2351 GTTACATATCAATATGTCTGACTTGAATGAACTGGAGCCACCATTCCTCGAGTAACT 2410
 QY 3236 CysSerProValSerCysGlyIysProGluSerProGluHisGlyPheValValGlySer 3255
 DB 2411 TGCAGTCCAGTTCTTGTGGGAACTGAAAGTCCAGAAACATGGATTGTGTGGGACAT 2470
 QY 3256 LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluIleuGluGly 3275
 DB 2471 AAATACACCTTTGAACACAAATTAATTAATCAAGTGAAGTGCCTGTATGAACAGAGGG 2530
 QY 3276 AsnArgGluArgValCysGluGluAsnArgGlnTrpSerGlyGlyValAlaIleCysIys 3295
 DB 2531 AACAGGAACGTGTCTGCCAGAGACAGACAGTGTGAGTGGAGGGTGGCATATGCAAA 2590
 QY 3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyValAlaAspIleGluAsnArg 3315
 DB 2591 GAGACAGAGTGTAACTCCACTGAATTTCTCATGTGAAAGCTGCATGTGAACAAACGG 2650
 QY 3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
 DB 2651 ACGACTGACCCCAACGTGTATATTCCTGCAACAGAGCTACAGCTTGAAGGGCCACT 2710
 QY 3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysIysProAsn 3355
 DB 2711 GAGGCACACTGCACAGAAATGGAACCTGAGCCACCAAGTCCCTCTGCAAAACCAAT 2770
 QY 3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluIleGluPheTyr 3375
 DB 2771 CCATGCCCTGTCTTGTGTGATTCGCCAGAAATGCTGCTGTGTGAAGAAAGATTAT 2830
 QY 3376 ValAspGluAsnValSerIleuValArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
 DB 2831 GTTGATCAGAAATGTCTCATCAATAATGAAGAGGATTTCTGTGCAAGGGCCACGGCATC 2890
 QY 3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaIysCysGluIleIleSer 3415
 DB 2891 ATTAACCTGCAACCCGACAGAGACTGACACAGCAAGCCGCAATGTGAATAATCTCA 2950
 QY 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
 DB 2951 TGTGTCTCACAGCTCAGTGAAGAAATGCAATTCCTGAGGCCGTATCAATATGGA 3010
 QY 3436 AspMetIleThrTyrSerCysTyrTrpSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
 DB 3011 GACATGATCACTCATGTATGATGAGATGATCAATGATGAGGGTTTCTGAGAGAGTGT 3070
 QY 3456 CysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
 DB 3071 TGTTTAATAAATGAACAATGACATGACCTCTATTTTGCAGAGCTGTCTGTGCATTTCCA 3130
 QY 3476 CysGluAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMet 3495
 DB 3131 TGTTCAGAAATGGGGCACTCTGCAACGCCCAATTCCTGTGTCTGTGTCCAGAGGGCTGAGT 3190
 QY 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysIleuAsnGlyIleArgCysVal 3515

Db 3191 GGGCCCTCTGTGAAGAACCAATGCAATCTTCCCTGCTACAGGAGCTGCTGTG 3250
 Oy 3516 AIAEPYVINGINCYASPCYSPROPGIYTPTHNGLYSEAPYCYSHIETHALAVL 3535
 Db 3251 GCCCTTACCACTGACGTGACCCGCTGCTGACGGGCTCTGCTGATACAGCTGTT 3310
 Oy 3536 CYGINSERPCYCLEUAMNGLYGLYLYSCYVALARGPROAMARGCYSHISCYLEU 3555
 Db 3311 TGCCAGTCTCCCTGCTTAATGCTGGAATAATGTAAGCAAAACGAGTCACTGCTT 3370
 Oy 3556 SERSERTPTHTNGLYHIAANCYSEARGLYRARGATHTNGLYPHE 3571
 Db 3371 TCTTCTTGACCGGACATMACTGTTCCAGAAAGAGAGACTGGGTTT 3418
 RESULT 18
 ABT07803
 ID ABT07803 standard; DNA; 3804 BP.
 AC ABT07803;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Novel human nucleic acid SEQ ID No 43.
 XX
 KW Immunostimulant; antirheumatic; antiarthritic; neuroprotective;
 KW antiallergic; antidibetic; antiaesthetic; antiinflammatory; nootropic;
 KW immunosuppressive; anticoagulant; thrombolytic; antithrombotic;
 KW cytostatic; nephrotoxic; antiparkinsonian; gynecological; virucide;
 KW antibacterial; antihistaminic; fungicide; HCPAT05; HMAE95; HTNEM01;
 KW immunodeficiency; autoimmune disorder; allergic reaction; cardiovascular;
 KW inflammatory condition; graft-versus-host disease; reproductive system;
 KW blood-related disorder; hyperproliferative; endocrine; neurological;
 KW respiratory; renal; infectious disease; gastrointestinal; gene therapy;
 KW neuronal growth; neuronal disorder; neuro-degenerative condition;
 KW keratinocyte growth; human; ds.
 KW
 OS Homo sapiens.
 XX
 PN US2002086330-A1.
 XX
 PD 04-JUL-2002.
 XX
 PF 17-JAN-2001; 2001US-00764893.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 26-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225470P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234979P.
 PR 29-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239352P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-024617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 XX
 PA (ROSE/) ROSEN C A.
 PA (RUBER/) RUBEN S M.
 PA (BARA/) BARASH S C.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI, 2002-665432/71.
 XX
 XX
 PT Novel polypeptide useful for diagnosis, prognosis, prevention, and
 PT treatment of immune, hyperproliferative, renal, respiratory,
 PT cardiovascular, reproductive, endocrine, gastrointestinal and
 PT neurological disorders.
 XX
 PS Disclosure, Page 229-231, 335pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising a sequence at
 CC least 90% identical to a full length protein sequence selected from 55
 CC sequences given in the specification such as a sequence of 163, 74 or 140
 CC amino acids fully defined in the specification, or the encoding sequence
 CC contained in 49 cDNA clones given in specification e.g. HCPAT05, HMAE95
 CC or HTNEM01. The protein and its encoding nucleic acid are useful for
 CC diagnosing a pathological condition or susceptibility to a pathological
 CC condition in a subject and for preventing, treating or ameliorating a
 CC medical condition. The protein, its encoding nucleic acid and an isolated
 CC antibody that can bind to the protein are useful in treating, preventing,
 CC diagnosing and/or prognosing immunodeficiencies, autoimmune disorders,
 CC allergic reactions and conditions, inflammatory conditions, graft-versus-
 CC host disease, blood-related disorders, hyperproliferative disorders,
 CC renal disorders, cardiovascular disorders, respiratory disorders,
 CC neurological disorders, endocrine disorders, reproductive system
 CC disorders, infectious diseases, and gastrointestinal disorders. The
 CC protein of the invention is useful to stimulate neuronal growth and to
 CC treat, prevent, and/or diagnose neuronal damage which occurs in certain
 CC neuronal disorders or neuro-degenerative conditions, for stimulating
 CC keratinocyte growth, to prevent hair loss, to modulate mammalian
 CC characteristic such as body height, weight, hair color, and to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components. The nucleic acid of the invention can be used in gene
 CC therapy. This polynucleotide sequence represents one of the novel nucleic
 CC acids of the invention
 XX
 SQ Sequence 3804 BP; 1025 A; 858 C; 915 G; 1006 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 8e-315 Length: 3804
 Score: 6452.00 Matches: 1135
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.91% Mismatches: 0
 Query Match: 32.30% Indels: 0

DB:	6	Gaps:	0
US-09-977-053-4 (1-3571) x AB07803 (1-3804)			
OY	2436	ValProValGluCyProGlnProGlnIleProAsnGlyIleIleAspValGlnGly	2455
DB	11	GTTCAGATGAAATATCCCAACTGAGAAATCCCAATGAAATCTTGAATGTGCAAGCC	70
OY	2456	LeuAlaIleuSerThrAlaLeuIleThrCysIleProGlyPheGluLeuValGlyAsn	2475
DB	71	CTTGCCATATCTCAGACAGGCTCTATATACCTGCAAGCCAGGCTTGAATGTGTGGAAAT	130
OY	2476	ThrThrThrIleuCyGlyGlyIleuAsnGlyIleIlePheGluGlyIleProThrCysIleVal	2495
DB	131	ACTACCAACCTTGTGTGAGAAATGTCACTGCTGGAGAGAAACCAACATGTAAACCC	190
OY	2496	IleGluCysIleuSerProGlyGluIleLeuAsnGlyIlePheSerIleThrAspLeuIle	2515
DB	191	ATTGAGTGCCTGAAACCCAGAGAAATTTGAAATGCAAAATCTCTTACAGGACCTTAC	250
OY	2516	TyrGlyGlnThrValThrIleCysAsnArgGlyPheArgLeuGluGlyProSerAla	2535
DB	251	TATGGAACAACCGTTACTTCTTGTGCAACGAGGCTTGGCTGGAAGTCCCAAGTCC	310
OY	2536	IleuThrCysIleuGlnThrGlyAspThrAspValAspAlaProSerCysAsnAlaIleHis	2555
DB	311	TTGACCGTTTAAAGACAGGTGATGGAGTGAATGCCCATCTTCGAAATGCCATCCAC	370
OY	2556	CysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAspIleThrCysIle	2575
DB	371	TGTGATTCCTCCCAACCCATTGAAATGTTTTGTGAAGGTGCAAGTAAAGCTTAAGT	430
OY	2576	AlaIleIleIleIleThrCysPheProGlyPheGlnValAlaGlyIleAlaIleThr	2595
DB	431	GCCATATATCATCTACAGTGTCTTCCCTGGGTTTCAAGGTGGCTGATGCCATGACAC	490
OY	2596	CysGluGluSerGlyTTPSerSerSerIleProThrCysMetProIleAspCysGlyIleu	2615
DB	491	TGTGAAGAGTCAGAGATGTCAAGTTCATCCCAATGTATGCCAATAGATCGAGCTTC	550
OY	2616	ProProIleIleAspPheGlyAspCysThrIleuIleuAspAspGlnGlyTTPPheGlu	2635
DB	551	CTCTCTCATATAGATTTTGAAGCTGATCTAACTCAAAAGATGACAGGATATTTTGA	610
OY	2636	GlnGluAspAspMetMetGluValProIleValThrProIleProProIleIleuGly	2655
DB	611	CAAGAAGACCAATGATGAGAAATTCATATGTGACTCTCACCTCTTATCATTTGGA	670
OY	2656	AlaValAlaIleThrTTPGluAsnThrIleuGluSerProIleThrIleSerSerAsnPhe	2675
DB	671	GCAGTGCTAAACCTGGAAATATCAAAAGAGTCTCTGCTACACATTCATCAAACTTT	730
OY	2676	IleuTyrGlyThrMetValSerIleThrCysAsnProGlyTyrGluLeuGlyAsnPro	2695
DB	731	CTGTATGTATCAATGTTTCAATACCTGTATCTCAAGATATGAACTTCTGGGAACTCT	790
OY	2696	ValIleuIleCysGlnGluAspGlyThrTTPAsnGlySerAlaProSerCysIleSerIle	2715
DB	791	GTCCTGATCTGCCAGAAATGAGAACTTGAAATGGCAGTGCACATCTCGATTTCAATT	850
OY	2716	GluCysAspIleuProThrAlaProGluAsnGlyPheIleuArgPheThrGluThrSerMet	2735
DB	851	GAATGATGACTTGCCTGCTCTCCGAAATGCGCTTTTGGTTTACAGACATCAGATG	910
OY	2736	GlySerAlaValGlnIleCysIleProGlyIleIleLeuAlaGlySerAspLeuArg	2755
DB	911	GGAAGTGTCTGCAATATCTGTAACCTGGAACATTCAGACAGCTCTGATTTAAG	970
OY	2756	LeuCysLeuGluAsnArgIleTTPSerGlyAlaSerProArgCysGluAlaIleSerCys	2775
DB	971	CTTGTGTATGAAATGAAAGTGAAGTGTGCTCCCAAGCTGTGTGAAGCATTTTATGC	1030
OY	2776	LybLybProAsnProValMetAsnGlySerIleuGlySerAsnIleThrIleuSer	2795
DB	1031	AAAAAGCCAAATCCAGTATGATGATCCAAAGAGACATCTACATCTTGAC	1090
OY	2796	ThrIleuTyrIleGluCysAspProGlyTyrValIleuAsnGlyThrGluAsnThrCys	2815
DB	1091	ACGTTGTATATGATGTGATCCCGGATATGTGCGAATGTGACATGAGAGGAAATGC	1150
OY	2816	GlnAspAspIleAsnTTPAspGluAspGluProIleCysIleProValAspCysSerSer	2835
DB	1151	CAGATGTCAAAAACCTGGAGTAGAGATGAGCCATTTCATCTCTGTGATCTGCACTTCA	1210
OY	2836	ProProValSerAlaAsnGlyGluValArgGlyAspGlyIleThrPheGlnIleGluIle	2855
DB	1211	CCCCAGTCTCAGCCAAATGSCCAGGTGAGAGACGAGTACATTCCTCAAAAAGATTT	1270
OY	2856	GluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAla	2875
DB	1271	GAATATCTTTCATATGAAAGGTTCTTCTTACAGGAGCCAGGAGTCCGGGTTGTCTGCC	1330
OY	2876	AsnGlySerTTPSerGlyAlaThrProAspCysValProValArgCysAlaThrProPro	2895
DB	1331	AATGGAAGTTGAGTGAAGCACCTCCAGCTGTGCTGTGACATGTGCCACCCGCCA	1390
OY	2896	GlnIleuAlaAsnGlyValThrGluGlyIleuAspTyrGlyPheMetIleGluValThrPhe	2915
DB	1391	CACCTGSCAAATGGGGTGAAGGCTTGACATATGCTTCTATAGAGATTAACATTC	1450
OY	2916	HisCysHisGluGlyTyrIleLeuHisGlyAlaProIleuLeuThrCysGlnSerAspGly	2935
DB	1451	CACGTGACGAGGGCTATCATCTTGACAGGTCTCCAAAACCTCAGTCAGTCAATGTC	1510
OY	2936	AsnTTPAspAlaGluIleProLeuCysIleProValAsnCysGlyProProGluAspLeu	2955
DB	1511	AACGTGATGACAGATTCCTCTGTAAACAGTCAACCTGAGACCTCTCAAAATCTT	1570
OY	2956	AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnThrIleCys	2975
DB	1571	GCCATGATTCCTCAATGATGTTTTCTTATATATGAGGAGCATATACAGTATGATGC	1630
OY	2976	PheProGlyTyrIleuIleGlyAsnSerSerArgAspCysLeuSerAsnGlySerTTP	2995
DB	1631	TTTCTGTATATAGCTTCCATGAAATTCATCAAGAAGTGTCTTCCATGTGCTCTGG	1690
OY	2996	SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly	3015
DB	1691	AGTGCAGCTCACTCTGCTGCTCTGAGATGTTCCACACCAAGTAAATGAATGGA	1750
OY	3016	ThrValAsnGlyThrAspPheAspCysGlyIleValAlaIleArgIleGlnCysPheIleGly	3035
DB	1751	ACTGTCAATGGAACAATTTTGAATGTGAAAGGCGAGCCGATTCAGTGTCAAAAGGC	1810
OY	3036	PheIleuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGlnTTPSerSerGly	3055
DB	1811	TTTCAAGCTCCATGAGACTTTTGAATATCACTGTAAAGCCATGTGCGATGAGACTTGG	1870
OY	3056	PheProHisCysGluHisThrSerCysGlySerIleuProMetIleProAsnAlaPheIle	3075
DB	1871	TTCCCCCACTGTGAACAACATCTTGTGTGTTCTCTTCCAAAGATACCAAAATCGTTTATC	1930
OY	3076	SerGluThrSerSerTTPValGluAsnValIleThrIleThrCysArgSerGlyTyrVal	3095
DB	1931	AGTGAACACACTCTTGAGAGAAATGTGATTAATCTTACAGCTGAGGTCTGGAATATTC	1990
OY	3096	IleGlnIleSerSerAspLeuIleCysThrGluIleGlyValTTPSerGlnProIlePro	3115
DB	1991	ATACAGGACGATTCAGATCTGATTTGTACAGAGAAAGGATGAGACACCTTATCA	2050
OY	3116	ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly	3135
DB	2051	GTCGTGAGGCTTGTCTGTGGGTGCCCAACCGTCTGTGCAATGCAAGTGGCACTGGA	2110
OY	3136	GluAlaHisThrIleThrIleuSerGluValIleuLeuArgCysLeuGluGlyTyrThrMetAsp	3155

Db 2111 GAGGACACACCTATGAAAGTAAAGTAACTCAGATGCTGAGAGGTATATACATGAT 2170
 Qy 3156 ThrAspThrAspThrPheThrCysGluLysAspGluArgTrpPheProGluIleSer 3175
 Db 2171 ACAGATACAGATACATTCACCTGTCAGAAAGATGTCGCTGAGCTCCCTGAGAAATCTCC 2230
 Qy 3176 CysSerProLysCysPheLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
 Db 2231 TGCAGCTCCAAAATAATGCTCTCCCGGAAAACATTAACATATATCTTGATCATGGGAGC 2290
 Qy 3196 AspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
 Db 2291 GATTCAGTGTGATATGAGCAAGTTCTGCTCATGTCAGAAAGGATATACCTTTGAGGGA 2350
 Qy 3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSer 3235
 Db 2351 GTTAACATATCAGATATGCTGAGCTGATGAACTCGGAGGCCACATTTCCGATGAAATCT 2410
 Qy 3236 CysSerProValSerCysGlyLysProGluSerProGluHisGlyPheValValGlySer 3255
 Db 2411 TGCAGCTCCAGTTTCTTGTGGGAAACCTGAAGTCCAGAAACATGATTTGTGGCAGT 2470
 Qy 3256 LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGly 3275
 Db 2471 AATATCACTTTGAAAGCAAAATTATTTATCAGTGTGAGCTGTGATATGAACTAGAGGCG 2530
 Qy 3276 AsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLys 3295
 Db 2531 AACAGGAAACGTGTCTGCGAGGAAACAGACGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2590
 Qy 3296 GluThrArgCysGluLysThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArg 3315
 Db 2591 GAGACCAAGGTGTGAATCTCCACTTGAATTTCTCAATGGGAAACCTGACATTTGAAACAGG 2650
 Qy 3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
 Db 2651 ACGACTGAGACCAACCTGGTATATTCCTGCAACAGAGGCTACAGTCTGTAAGGCGCATCT 2710
 Qy 3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn 3355
 Db 2711 GAGGCACTGCAACAAATGGAACCTGAGCCACCCGCTCTCTGCAACCAAAAT 2770
 Qy 3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr 3375
 Db 2771 CCATGCCCTGTTCTTTGTGATTCGCGAAGATGCTGTGTCGTAAGAAAGAGTTTAT 2830
 Qy 3376 ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
 Db 2831 GTTGAATCAGAAATGTGTCCATCAATATAGGGAAGTTTCTGTGCAAGGCGCCACGCGATC 2890
 Qy 3396 IleThrCysAsnProAspGluThrTrpTrpGlnThrSerAlaLysCysGluLysIleSer 3415
 Db 2891 ATTACTGTGACCCCGACGAGAGCTGAGACAGACAGAGCGCCAAATGTGAAAAATCTCA 2950
 Qy 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
 Db 2951 TGTGTGCCACCGCTCACCTAGAAATGCAATTGCTCGAGGCGCTATATCATATGCA 3010
 Qy 3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
 Db 3011 GACATGATCATCTACTCATGTTTACAGTGTATCATGATGTGAGCGGTTTCCGAGAGAGTGT 3070
 Qy 3456 CysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
 Db 3071 TGTTTGAAAAATGAAACATGAGACATCACTCTATTTTGAGAGCTGTCTGTCAATTTCCA 3130
 Qy 3476 CysGlnAsnGlyValIleCysGlnArgProAsnAlaCysSerCysArgProGluGlyTrpMet 3495
 Db 3131 TGTCAGAATGGGGGAGCATCTGCCAAGCCCAAAATGTTGTTCTGTCTCAAGAGGCGTGAATG 3190
 Qy 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyValArgCysVal 3515
 Db 3191 GGGCGCTCTGTGAAAGAACCAATGTGCAATTTCTCTGCTGAACGAGGTCGCTGTGTG 3250

Qy 3516 AlaProTyrGlnCysAspCysProProGlyTyrTrpThrGlySerArgCysHisThrAlaVal 3535
 Db 3251 GCCCTTACAGAGTACATCCCGCTGCTGAGACGGGAGCTGCTGTATACAGCTGT 3310
 Qy 3536 CysGlnSerProCysLeuAsnGlyValCysValArgProAsnArgCysHisCysLeu 3555
 Db 3311 TGCAGTCTCCCTGCTTAATGTGAAATGTGTAAAGCAAAACCATGTCTCACTGCTT 3370
 Qy 3556 SerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe 3571
 Db 3371 TCTTCTTGACGGGACATTAATCTTTCAGGAAAAGAGAGCTGGGTTT 3418
 RESULT 19
 ABV83885
 ID ABV83885 standard; cDNA; 3804 BP.
 AC ABV83885;
 XX
 DT 09-DEC-2002 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 214.
 XX
 KW Human; noctropic; neuroprotective; cytostatic; dermatological; virocidic;
 KW Immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
 KW antiparkinsonian; antistroke; antianemic; antidiabetic; cancer;
 KW antihypertensive; hepatoprotective; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
 KW antiparkinsonian; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotoxic; gene therapy; vaccine;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002090672-A1.
 PD 11-JUL-2002.
 XX
 XX
 PE 17-JAN-2001; 2001US-00764853.
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 11-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225477P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 30-AUG-2000; 2000US-022824P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 21-SEP-2000; 2000US-0234213P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234977P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236377P.

1211 CCCCAGTCTCAGCCAAATGCGCAGGTGAGAGAGACGATGACATTTCCAAAAAGAGATT 1270
QY
2856 GIUTYRThrCysAsnGluGlyPheLeuLeuGluGlyValArgSerArgValCysLeuAla 2875
DB
1271 GAATACATCTTGCAATGAGGGTTCTTGCTTGAGGGAGCCAGAGGTGGGTTTGTCTTGC 1330
QY
2876 AsnGlySerTrpSerGlyValThrProAspCysValProValArgCysValThrPro 2895
DB
1331 AATGGAAAGTTGAGGTGAGCCACTCCGACTGTGCTCTGAGATGTGCAACCCGCCCA 1390
QY
2896 GlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetCysGlyValThrPhe 2915
DB
1391 CAACGGCCCAATGGGGTGAAGGAGCCCTGAGCTATGGCTTCATGAAAGAAATGAACATTC 1450
QY
2916 HisCysHisGlyGlyTyrTlleLeuHisGlyValProLeuLeuThrCysGlnSerAspGly 2935
DB
1451 CACTGTCAGAGGGCTACATCTTGCAACGGTCTCCAAAACCTGACTGCTGAGATGGC 1510
QY
2936 AantTrpAspAlaGluIleProLeuCysLeuProValAsnCysGlyProProGluAspLeu 2955
DB
1511 AACTGGAGTGCAGAGATTCTCTCTGTAAACAGTCAACGTGAGACTCTCGAAGATCTT 1570
QY
2956 AlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCys 2975
DB
1571 GCCCAATGTTTCCCTAATGGTTTTCCTTATTCATGGGGGCCATATACAGTACAGTGC 1630
QY
2976 PheProGlyTyrTyrLeuHisGlyAsnSerArgArgCysLeuSerAsnGlySerTrp 2995
DB
1631 TTTCCGAGTTATAGCTCCATGGAAATTCATCAAGAAAGTCCCTTCATAGGCTCTG 1690
QY
2996 SerGlySerSerProSerCysLeuProCysArgCysSerThrProValIleGluTyrGly 3015
DB
1691 AGTGGAGCTCAACCTTCCTGCTGCTGAGATGTTCCACACAGTAATTTGAATTTGA 1750
QY
3016 ThrValAsnGlyTyrTrpAspPheAspCysGlyValAlaIleArgIleGlnCysPheLeuGly 3035
DB
1751 ACTGTCAATGGGACAAATTTGACTGTGAAAGGAGCCGAGATTCAGTCTTCAAGGC 1810
QY
3036 PheLeuLeuLeuGlyLeuSerGluIleThrCysGlyValAlaAspGlyGlnTrpSerGly 3055
DB
1811 TTCAGCTCTAGAGCTTTCTGAATCACCCTTGGAAGCCGATGCCAGTGGACTCTGGG 1870
QY
3056 PheProHisCysGlyHisIleThrSerCysGlySerLeuProMetIleProAsnAlaPheIle 3075
DB
1871 TTCCTCCACCTGTAACAACCTCTGTGGTCTCTTCCAAATGATACAAATGGGTCATC 1930
QY
3076 SerGluThrSerSerTrpIlySerGluAsnValIleThrTyrSerCysArgSerGlyTyrVal 3095
DB
1931 AGTGAGACCAAGCTCTTGAAAGGAAATGTGATTAACAGCTGACAGGTGGAATATGTC 1990
QY
3096 IleGlnGlySerSerAspLeuIleCysThrGlyLeuGlyValTrpSerGlnProTyrPro 3115
DB
1991 ATACAAGGCAAGTTCAATCTGATTTGTACAGAAAGGGGTATGAGCCGCTTATCCA 2050
QY
3116 ValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGly 3135
DB
2051 GCTGTGAGACCCCTTGTCTGTGGGTGCCCAACCTGTCTGCGCAATGAGGAGCACTGGA 2110
QY
3136 GluAlaHisIleThrTyrGluSerGlyValIleLeuLeuArgCysLeuGluGlyTyrThrMetAsp 3155
DB
2111 GAGGCAACAACCTATAGAAAGTGAAGTGAACCTAGAGTGTGAAAGGTTATACAGATGAT 2170
QY
3156 ThrAspTrpAspThrPheThrCysGlnIleAspGlyValArgTrpPheProGluArgIleSer 3175
DB
2171 ACAGATACAGATACATTCACCTGTCAAGAAAGTGTGCTGCTCCGTGAGAGATCTCC 2230
QY
3176 CysSerProLeuLeuCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsp 3195
DB
2231 TCCAGGCTCTAAAGAAATGTCTCTCCCGGAAACATTAACATATACCTGACAGGGGAC 2290
QY
3196 AspPheSerValAsnArgIleValSerValSerCysAlaGluGlyTyrThrPheGluGly 3215
DB
2291 GATTCAGTGTGATAGGCAAGTTTCTGTGTCATGTGCAGAAAGGATACCTTTAGAGGA 2350

QY
3216 ValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSer 3235
DB
2351 GTTAACATATCAAGTATGTCAAGCTTGATGGAACCTGGGAGCCACATCTCCGATGAAATCT 2410
QY
3236 CysSerProValSerCysGlyLeuProGluSerProGluHisGlyPheValValGlySer 3255
DB
2411 TGCAGTCCAGTTTCTGTGGGAAACCTGAAAGTCCAGAAATGGAATTTGTGGTGGCAGT 2470
QY
3256 LysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGluTyrTyrGluLeuGluGly 3275
DB
2471 AATTAACCTTTAAAGCACAATTAATTAATCAAGTGAAGCTGAGCTTGAATGAGGGG 2530
QY
3276 AsnArgGluArgValCysGlnGluAsnArgIleTrpSerGlyValAlaIleCysLeu 3295
DB
2531 AACAGGGAAGTGTCCGCGAGAAACAGACAGTGAAGTGAAGGGGTGCAATATGCAAA 2590
QY
3296 GluThrArgCysGluThrProLeuGluPheLeuAsnGlyValValAlaAspIleGluAsnArg 3315
DB
2591 GAGACCAAGTGTGAACCTCACTGAATTTCTCAATGGGAAAGCTGACATTTGAATAACAG 2650
QY
3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
DB
2651 ACAGCTGAGCCCAAGTGTGATATTCCTGCAACAGAGCTACAGTCTGAAAGGCCATCT 2710
QY
3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLeuProAsn 3355
DB
2711 GAGGCACTGCAAGAAATGGAACCTGAGCCACCCAGTCCCTCTCCAAACCAAAAT 2770
QY
3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGlyLeuGluPheTyr 3375
DB
2771 CCAATGCCCTGTCTTGTGTGATTTCCGAAAGTCTGTGCTGTGAAAGAGTTTAT 2830
QY
3376 ValAspGluAsnValSerIleLeuCysArgGluGlyPheLeuLeuGlnIleHisGlyIle 3395
DB
2831 GTTGATCAGATGTGTCCATCAATGTAGGAAAGTTTCTGCTGCAAGGCCCAAGCATC 2890
QY
3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLeuCysGlyLeuValSer 3415
DB
2891 ATTACCTGCAACCCCAAGCAAGTGGACACAGACAAAGCCCAAAATGTGAAAAATCTCA 2950
QY
3416 CysGlyProProAlaHisValGluAsnAlaIleAlaAspGlyValHisGlyGlnTyrGly 3435
DB
2951 TGTGTCTCAACACTCACTGTAAGAAATGTGATGTCTGAGGCGTACATTCATATGGA 3010
QY
3436 AspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
DB
3011 GACATGATCACTACATCTTACAGTATACATGTGAGGGTTTCCGAGGAGTGT 3070
QY
3456 CysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCysArgPhePro 3475
DB
3071 TGTTTGAAATGTGAACATGACATCACTCCATATTTGCAAGAGCTGTCTGTCAATTTCCA 3130
QY
3476 CysGluAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMet 3495
DB
3131 TGTCAAAATGGGGCAATCTGCAAGCCCAAAAGCTGTGTCTGTCCAGAGGCTGATG 3190
QY
3496 GlyArgLeuCysGluGluProIleCysHisIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
DB
3191 GGGGCGCTCTGTGAAGAACCAATCTGATCTTCCCTGTCTGAAACGAGGTGCTGTGTG 3250
QY
3516 AlaProTyrGlnCysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
DB
3251 GCCCTTACCAAGTGTATCCCGCGCTGAGAGGGGAGTCTGCTGTCAACAGCTGTT 3310
QY
3536 CysGlnSerProCysLeuAsnGlyGlyLeuCysValArgProAsnArgCysHisCysLeu 3555
DB
3311 TGCAGTCTCTCTGTAAATGTGAAATGTGTAAGACCAACCAATGTCACTGTCTT 3370
QY
3556 SerSerTrpThrGlyHisAsnCysSerArgLeuArgArgThrGlyPhe 3571
DB
3371 TCTTCTTGACGGGACATTACTGTCCAGGAAAGAGAGAGCTGGT 3418

RESULT 20
AAH75787
ID AAH75787 standard; cDNA; 3696 BP.
AC AAH75787;
XX 16-OCT-2001 (first entry)
XX Receptor 222 coding sequence.
XX Receptor 222: cytostatic; anti-HIV; cancer; HIV infection; ss.
XX Unidentified.
OS
XX Key Location/Qualifiers
XX FH 402..3342
XX PT CDS /*tag= a
XX PT /product= "Receptor 222"
XX
XX CN1296952-A.
XX PD 30-MAY-2001.
XX 23-NOV-1999; 99CN-00124086.
XX PR 23-NOV-1999; 99CN-00124086.
XX (SHAN-) SHANGHAI BORONG GENE DEV CO LTD.
XX Mao Y, Xie Y;
XX MPI; 2001-483894/53.
XX P-PSDB; AAG66398.
XX New polypeptide-complement receptor 222 for treating diseases, such as,
XX cancer and human immunodeficiency virus infection.
XX
XX Claim 6; Page 19-21 (Disclosure); 30pp; Chinese.
XX
XX The present sequence is the coding sequence for receptor 222. Receptor
XX CC 222 and its coding sequence are useful in treating diseases e.g. cancer
XX and HIV infection
XX
XX SQ Sequence 3696 BP; 977 A; 837 C; 895 G; 987 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 9, 33e-309 Length: 3696
Score: 6331.00 Matches: 1113
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.70% Indels: 0
DB: 4 Gaps: 0
US-09-977-053-4 (1-3571) x AAH75787 (1-3696)
QY 2459 LeuSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThr 2478
DB 1 CTCAGACAGCTCTCTATACCTGCAGCCAGGCTTGAATTGGTGGAAATATCTACACC 60
QY 2479 LeuGlyGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 2498
DB 61 CTTTGTGGAGAAATGTCATGCTGGCTGGAGAAAACCAACATGTAAAGCCATTGAGTGC 120
QY 2499 LeuLysProLysGluLeuLeuGluGluGluGluGluGluGluGluGluGluGluGlu 2518
DB 121 CTGAAACCAAGAGATTTTGAATGCAATTCTCTTACCGGACCTTACCTATGAGACG 180
QY 2519 ThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCys 2538
DB 181 ACCGTTACTACTCTTGCAACCGAGGCTTCGCTCGAAGGTCCAGTGCCTTGACCTGT 240
QY 2539 LeuGluThrGlyAspThrAspValAspAlaProSerCysAsnAlaLeuHisCysAspSer 2558

DB 241 TTAGACACAGTGAATTGGAGTGTAGATGCCCATCTTGCAATGCCATCCATGATTC 300
QY 2559 ProGluProIleGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 2578
DB 301 CCACACCCATTGAAATGTTTGTAGAAAGTGCAGATTTACACTATGGGCCATTATC 360
QY 2579 IleTyrSerCysPheProGlyPheGluValAlaGlyHisAlaMetGluThrCysGluGlu 2598
DB 361 ATCTACAGTGTCTCCCTGGGTTTCAGGTGCTGTGTATGTCATGCGAGACCTGTGAAG 420
QY 2599 SerGlyThrSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHis 2618
DB 421 TCAGATGTCAGATTCATCCCAACATGATGCAATGATGATGATGATGATGATGATGAT 480
QY 2619 IleAspPheGlyAspCysThrLysLeuLysAspAspGluGlyTyrPheGluGluGluAsp 2638
DB 481 ATGATTTTGGAGACTGTACTTAACTCAAGATGACCAAGGATTTTGGACAAAGAGAC 540
QY 2639 AspMetMetGluValProTyrValThrProHisProProTyrHisLeuGlyValAla 2658
DB 541 GACATGATGGAAGTTCATATGTAATCTCACCCCTCTTATCATTTGGAGCAGTGCT 600
QY 2659 LysThrThrGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 2678
DB 601 AAAACCTGGAGAAATATCAAGAGAGTCTCTGCTACACATTCATCAACTTCTGTATGCT 660
QY 2679 ThrMetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGluGluGluGluGlu 2698
DB 661 ACCATGTTTCATACACTGTATATCCAGATATGAATCTTCGGGAAACCTGTCTGATC 720
QY 2699 CysGluGluAspGlyThrThrAsnGlySerAlaProSerCysIleSerIleGluCysAsp 2718
DB 721 TGCAGAGAAATGGAATTCGAAATGGCAGTGCAACATCTGCATTTCAATTTGAATGAC 780
QY 2719 LeuProThrAlaProGluLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 2738
DB 781 TTGCTACCTGCTCTGAAATGCTTTTTCGCTTTTACAGAGACTAGACAGAGAGTCT 840
QY 2739 ValGluTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeu 2758
DB 841 GTGAGATATGCTGTAAACCTGACACATCTGACAGGCTGTGACTTAAGCTTTGTCTA 900
QY 2759 GluAsnArgLysThrSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysPro 2778
DB 901 GAGAAATAGAAAGTGTGTGCTCCCAACCTGTGAAACCATTTATGAGAAAACCA 960
QY 2779 AsnProValMetAsnGlySerIleLysGlySerAsnThrThrTyrLeuSerThrLeuTyr 2798
DB 961 AATTCAGTCAATGATGATCATCAAGAGAGCACTACACATCTGAGCAGCTGTTC 1020
QY 2799 TyrGluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGluAspAsp 2818
DB 1021 TATGAGTGTGACCCCGGATATGTGTGAATGGCATGAGAGAGAAACATGCAAGATGAC 1080
QY 2819 LysAsnThrAspGluAspGluProIleCysIleProValAspCysSerSerProProVal 2838
DB 1081 AAAAATCGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 2839 SerAlaAsnGlyGluValArgGlyAspGluTyrThrPheGluGluGluGluGluGluGlu 2858
DB 1141 TCAGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 2859 CysAsnGluGlyPheLeuLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 2878
DB 1201 TGCATGAGAGGTTCTTGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 2879 ThrSerGlyAlaThrProAspCysValProValArgCysAlaThrProProGluLeuAla 2898
DB 1261 TGAAGTGAAGCCTCCCACTGTGTGCTGTGATGATGATGATGATGATGATGATGATGAT 1320
QY 2899 AsnGlyValThrGluGluLeuAspTyrGlyPheMetLysGluValThrPheHisCysHis 2918
DB 1321 AATGGGTGACGAAAGGCTGAGCTATGCTTATGAGAGATTAATTCATCTCATCTCAT 1380

QY 2919 GtuglytYrTlletleuhtsglyalaProlyleuThrcYsglInseraspGlyAsnTTPAP 2938
 DB 1381 GAGGCTACATCTTTCACCGCTCTCCAAAATCACTGCTCAGTCAGATGCACTGGAT 1440
 QY 2939 AAggluIleProleuCyalyaProValaAnCyeglyProProgluApleuAhiAgly 2958
 DB 1441 GAGAGATATCTCTCTGTAACCAAGTCAGCTGAGCTCTGAAATATCTTCGCTCAT 1500
 QY 2959 PheProaAnGlyPheSerPheIleHISglYglYHISglYrTnCySerPheProGly 2978
 DB 1501 TTCCTTAAGTGTTCCTTTCCTTATTCATGAGGAGCATATACAGTATAGGCTTCG 1560
 QY 2979 TTrlyleuHISglYAsnSerSerArgArgCyalyeSerAsnGlySerTTPSerGlySer 2998
 DB 1561 TATTAAGCTCATAGGAATTCATCAACAGAGGCTCTCCAAATGCTCTGAGATGGCAG 1620
 QY 2999 SerProSerCyalyeProCyalyaArgCySerThrcProValIlegluYrTnYrThrcValaAn 3018
 DB 1621 TCACCTTCCTGCTGCTTCGAGATGTTCCACACAGTAATGAAATATGAACTGTCAT 1680
 QY 3019 GtYrThrcPheAspCyeglyYrYsaIaAlaArgIlegluCySerPheIleGlyYrThrc 3038
 DB 1681 GGGACAGATTTTGACTGTGAAAGGCAAGCCGATTCAGTCTTCAAGGCTTCAAGCTC 1740
 QY 3039 IeuglyleuSerGluIleThrcYsglyuaIaAspGlygluIleThrcSerGlyPheProHIS 3058
 DB 1741 CTAGACTTCTGAAATACCTCCAGAGCCGATGGCAGTGGAGCTCTGGATCCCCCAG 1800
 QY 3059 CyeglyuHISerThrcYsglySerleuProMetIleProAsnAlaPheIleSerGluThrc 3078
 DB 1801 TGTGAACACACTTCTGTGTGTCTCTTCCAAATGATACCAATCCGTTCATCAGTGAACC 1860
 QY 3079 SerSerTTPlygluAsnValIleThrcYrSerCyalyaSerGlyYrYrYrYrYrYrYr 3098
 DB 1861 AGCTCTTGAGAAATGTGATACCTTACAGCTGAGGCTGAGATATGATACATACAGG 1920
 QY 3099 SerSerAspLeuIleCySerThrcGlyYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3118
 DB 1921 AGTTCAAGATCTGATTTGTACAGAGAAAGGGGTATGAGCCAGCTTATCAGCTGTGAG 1980
 QY 3119 ProleuSerCyeglySerProProSerValaIaAsnAlaValaIaThrcIlegluAhi 3138
 DB 1981 CCTTGTCTGTGAGGCCACCGCTGTCTGCGCAATGCAAGTGCACCTGAGGAGCAC 2040
 QY 3139 ThrYr 3158
 DB 2041 ACCTATGAAAGTGAAGTGAAGTCAAGATGTCTGAAAGTTATTCAGATGATACAGAT 2100
 QY 3159 AspThrcPheThrcYsglyuIleYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3178
 DB 2101 GATACATTCACCTGTCAGAAAGATGCTCCCTGCTCCCTGAGAGATCTCTGCAATCC 2160
 QY 3179 LyalyeCyalyeProleuProgluAnIleThrcHISglYleuValaHISglYAspPheSer 3198
 DB 2161 AAAAAATGCTCTCCCGGAAAAACATACACATATCTGTACATGAGGAGCATTTCAAG 2220
 QY 3199 ValaAsnArgIleValaSerValaSerCyalyaIlegluYrTnThrcPhegluIleValaAnIle 3218
 DB 2221 GTGAATAGGCAAGTTCGTGTATGTGCAAGAGGATATACCTTGAAGGAGTTAACATA 2280
 QY 3219 SerValaCyeglyuIleuAspGlyThrcTTPgluProProPheSerAspGlySerCyalyePro 3238
 DB 2281 TCAGATATGCACTGTAAGTGAAGCTGAGGACCACTTCTCCGATGAATCTTGCATCCA 2340
 QY 3239 ValSerCyeglyYr 3258
 DB 2341 GTTCTTGTGGGAAACCTGAAAGTCCAGAACATGATTTGTGTGCTGCGATTAATACACC 2400
 QY 3259 PhegluSerThrcIleIleYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3278
 DB 2401 TTGGAAGCAATATTTATCAGTGTGAGCTTGGCTATGAACTTAAGGGGAAACAGGAA 2460

QY 3279 ArgValaCyeglyuIleuAsnArgIleThrcSerGlyYrYrYrYrYrYrYrYrYrYrYr 3298
 DB 2461 CGTGTCTGCAGAGAAACAGACAGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
 QY 3299 CyeglyuThrcProleuGluPheleuAsnGlyYrYrYrYrYrYrYrYrYrYrYrYr 3318
 DB 2521 TGTGAACCTCCACTTGAATTTTCAATGGAAGGAGTGAACATTTGAACACAGAGACTG 2580
 QY 3319 ProAsnValaIleYrSerCyalyaAsnArgGlyYrYrSerleuGluIleProSerGluAhi 3338
 DB 2581 CCCAGCTGTATATTCCTGCAACAGAGGCTACAGCTTGAAGGAGCATGAGGAGCAC 2640
 QY 3339 CyethrcIleuAsnGlyThrcTTPSerHISProValaProleuCyalyaProAsnProCy 3358
 DB 2641 TGCACAGAAATGGAACCTGAGCCACCAAGTCCCTCTGCAAAACCAATTCATGCTCT 2700
 QY 3359 ValProPheValIleProgluAsnAlaIleuSerGlyYrYrYrYrYrYrYrYrYrYr 3378
 DB 2701 GTTCCTTTGTGATTCCTGCAAGATGCTGTGTCTGAAAGGAGTTTATGTTATCAG 2760
 QY 3379 AsnValSerIleYrCyalyaArgGlyPheleuGluIleYrYrYrYrYrYrYrYrYrYr 3398
 DB 2761 AATGTCTCATCAAAATGTAAGGAGGTTTCTGTGCAAGGAGGAGGAGGAGGAGGAGG 2820
 QY 3399 AsnProAspGluThrcTTPThrcIleThrcSerAlaYrYrYrYrYrYrYrYrYrYr 3418
 DB 2821 AACCCGACAGAGCAGTGAACACAGACAGAGCCCAATGTGAATAATCTCATGTGTCCA 2880
 QY 3419 ProAlaHISValaGluAsnAlaIleAlaArgGlyYrYrYrYrYrYrYrYrYrYrYrYr 3438
 DB 2881 CCAGCTCAGTGAATAATGCAATGTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2940
 QY 3439 ThrYrSerCyalyeSerGlyYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3458
 DB 2941 ACTTACTCATGTTACAGTGAATACATGTTGAGAGGTTCTCTGAGAGGTTGTTAGAA 3000
 QY 3459 AsnGlyThrcTTPThrcSerProProIleCyalyaAlaValaCyalyaPheProCyeglyuAn 3478
 DB 3001 AATGGAACATGACATCACTCTTATTCAGAGCTGTCTGTGATTCATTCATTCAGAA 3060
 QY 3479 GtYr 3498
 DB 3061 GGGGAGATCTGCACAGCCCAATGCTTGTCTGTGCAAGAGGAGGAGGAGGAGGAGG 3120
 QY 3499 CyeglyuGluProIleYr 3518
 DB 3121 TGTGAAGAACATTCGATTCCTCCCTGTCTGAACGAGAGTGTCTGTGAGCCCTTAC 3180
 QY 3519 GlnCyalyaAspCyalyeProProgluYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3538
 DB 3181 CAGTGTGACTGCCCGCTGCTGAGACGGGAGTCTCGCTGTATACAGCTGTTCAGCTCT 3240
 QY 3539 ProCyalyeAsnGlyGlyYr 3558
 DB 3241 CCTGCTTAAATGAGTGAAGATGTGAAGACCAACAGATGATCTTCTTCTG 3300
 QY 3559 ThrGlyHISAsnCyalyeSerArgYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 3571
 DB 3301 ACCGAGCATACCTTTCAGAGAAAGAGAGACTGGTTT 3339

RESULT 21
 AEN93421
 ID AEN93421 standard; cDNA; 3706 BP.
 XX
 AEN93421;
 AC
 XX 23-UTL-2002 (first entry)
 XX
 XX Human gene GS9663 coding sequence from chromosome 9q31-34 #6.
 XX Human gene; chromosome 9q31-34; lipid protein metabolism disorder;
 KW cholesterol transport disorder; ss.
 XX

OS Homo sapiens.
 XX
 PN MO200071710-A2.
 XX
 PD 30-NOV-2000.
 XX
 PF 25-MAY-2000; 2000MO-FR001426.
 XX
 PR 25-MAY-1999; 99FR-00006587.
 PR 16-JUN-1999; 99US-0139450P.
 XX
 XX (AVERT) AVENTIS PHARMA SA.
 XX
 XI Denefle P, Rosier-Montus M, Arnould-Reguigne I, Prades C;
 PI Clepet C;
 XX
 DR MPI; 2001-025161/03.
 DR P-PSDB; ABR37987.
 XX
 PT New nucleic acid derived from human chromosome 9, used e.g. for diagnosis
 PT and drug screening, derived from genes implicated in disorders of
 PT lipoprotein metabolism.
 XX
 PS Claim 1; Page 197-199; 269pp; French.
 XX
 CC The present sequence is the coding sequence for a human gene from
 CC chromosome 9q31-34. This sequence is likely to be involved in diseases of
 CC plasmatic lipoprotein metabolism, e.g. the reverse transport of
 CC cholesterol
 XX
 SQ Sequence 3706 BP; 974 A; 837 C; 892 G; 995 T; 0 U; 8 Other;
 XX
 Alignment Scores:
 Pred. No.: 2 47e-294 Length: 3706
 Score: 6044.00 Matches: 1063
 Percent Similarity: 99.72% Conservative: 0
 Best Local Similarity: 99.72% Mismatches: 3
 Query Match: 30.26% Indels: 0
 DB: 4 Gaps: 0
 US-09-977-053-4 (1-3571) x ABRN3421 (1-3706)

QY 2666 GluSerProAlaThrHisSerSerAspPheLeuTyrGlnMetValSerTyrThrCys 2685
 DB 482 GAGTCTCTGTGTCACATTCATCAAACTTCTGTATGGTACCAAGTTCATACACCTGT 541
 QY 2686 AspProGlyTyrGlnLeuLeuGlyAsnProValLeuIleCysGlnGluAsnGlyThrTrp 2705
 DB 542 AAMCAGATATGAACTTCGGGGAAACCTGCTGATCTTCCAGAAAGATGAACTTGG 601
 QY 2706 AsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsn 2725
 DB 602 AATGGACGTGACCAATCTGATTCATTAATGATCACTTGCTTCTGCTTCTGAAAT 661
 QY 2726 GlyPheLeuArgPheThrGlnThrSerMetCysSerAlaValGlnTyrSerCysIlePro 2745
 DB 662 GCGTTTGTGGCTTTTACAGAGACTAGCATGGGAAGTGTGTCAGTATAGCTGTAACCT 721
 QY 2746 GlyHisIleLeuAlaGlySerAspLeuArgLeuGlyCysLeuGluAsnArgIleTrpSerGly 2765
 DB 722 GGACACATTCATGAGGCTCTGACTTAAGCTTGTCTAGAGAAATGAAAGTGAAGTGTG 781
 QY 2766 AlaSerProArgCysGluAlaIleSerCysIleSerProAsnProValMetAsnGlySer 2785
 DB 782 GCTTCCCAACCTGTGAAGCATTTCATGCAAAAGCCAAATCCAGTCAATGATGATCC 841
 QY 2786 IleGlySerAlaSerTyrThrTyrLeuSerThrLeuTyrGlnCysAspProGlyTyr 2805
 DB 842 ATCAAGAGAACATCACTACATACCTGAGCAGTGTGATGATGAGATCCCGGATAT 901
 QY 2806 ValLeuAsnGlyThrGlnArgArgThrCysGlnAspAspIleAsnTrpAsnGluAsnGlu 2825
 DB 902 GTGCTGAATGACATGAGAGAGAAACATGCGAGATGACAAACTGGAAGATGAGATAG 961
 QY 2826 ProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyValAlaArg 2845
 DB 962 CCCATTTCATCTCTGTGACTGCAATTCACCCCAAGTCAAGCCAAAGGCAAGTGA 1021
 QY 2846 GlyAspGluTyrThrPheGlnLeuValGlnIleGlnTyrThrCysAsnGlyGlyPheLeuLeu 2865
 DB 1022 GGAGACAGATACATTCAAAAGATGAATACCTTCATTAAGAGGGTTCCTGCTT 1081
 QY 2866 GluGlyValaArgSerArgValCysLeuAlaAsnGlySerIleTrpSerGlyValaThrProAsp 2885
 DB 1082 GAGGAGCCAGAGATGAGGTTTGTCTTGCATGAAGTGAAGTGAAGGCACTCCGAC 1141
 QY 2886 CysValaProValArgCysAlaThrProProGlnLeuAlaAsnGlyValaThrGlnGlyLeu 2905
 DB 1142 TGTGTGCTGTGATGTGCCACCCGCCACACTGGCCATGGGGTGAACGAGGCTTG 1201
 QY 2906 AspTyrGlyPheMetIleGlyValaThrPheHisCysHisGlnGlyTyrIleLeuHisGly 2925
 DB 1202 GACTATGGCTTCATGAAGAGATCACTTCCACTGATGAGGGCTACATCTTGCAACGCT 1261
 QY 2926 AlaProIleLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlnIleProLeuCysIle 2945
 DB 1262 GCTTCAAACTCAACCTTCACTGATGAGGCAACGAGATGAGAGATTCCTCTGTAAA 1321
 QY 2946 ProValAsnCysGlyProProGlnAspLeuAlaHisGlyPheProAsnGlyPheSerPhe 2965
 DB 1322 CCACTCAACTGTGACCTCTGGAAGATCTTGCCATGCTTCCCTAATGATTTTCCCTT 1381
 QY 2966 IleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrIleLeuHisGlyIleAsnSer 2985
 DB 1382 ATTCAATGGGGCCATATACAGTATCACTGCTTCTCTGTATTAAGCTCCATGAGAAATCCA 1441
 QY 2986 SerPheArgCysLeuSerAlaAsnGlySerTrpSerGlySerSerProSerCysLeuProCys 3005
 DB 1442 TCAAGAAAGTGGCTCTCAATAGCTCTGAGAGTGGCAGCTCACTTCCTGCTGCTG 1501
 QY 3006 ArgCysSerThrProValIleGlnTyrGlyThrValAsnGlyThrAspPheAspCysGly 3025
 DB 1502 AGATGTTCCACACCAAGTATTTGAATATGAACTGTCAATGGGACAGATTTTGAAGTGA 1561

QY	3026	LYSLALALARRGLLEGINCYSPHELYSGIYPHELYSLLEUENGLYLEUSERGILIETHR	3045
DB	1562	AAGCAACCCGGATTCAGTGGCTTAAAGCTTCAGACTCTTAGACCTTTCTGAATACCC	1621
QY	3046	CYSGIUNALAAEPGIYGINTRPSESGIYPHEPROHISCYSGIUNISHTRSERCYSGIY	3065
DB	1622	TGTAAACCCGATGSCAGTGGAGCTCTGGGTCCCACTGTGAACACATTCCTGTGGT	1681
QY	3066	SERLEUPROHETLEPROAHNALPHEIIESERGILNHSERSETTPLYSGIUNSERVAL	3085
DB	1682	TCTCTCCAAATGAAACCAAAATGCCATTCAGTAGACCAAGCTCTTGAAGAAAAGTG	1741
QY	3086	ILETHRYSERCYARSERGIYTYVALIILEGINGLYSERSEAPLEULIECYETHR	3105
DB	1742	ATAACTTACACTCAAGCTCTGGATTAATGATACAAAGCAAGTTCAAGATTTGTACA	1801
QY	3106	GIUNYSGIYVALTRPSEGINPROTYPROVALCYSGIUNPROLEUSERCYSGIYSEPRO	3125
DB	1802	GAGAAAGGGATATGAGACCACTTATCCAGTGTGTGAGCCCTGTGCTGTGGGTCCCA	1861
QY	3126	PROSERVALAIAANALAVALLALETHGYGLUALAHISTHRYRGUSERGIUNALYIS	3145
DB	1862	CCGCTGTGCGCAATGCACTGGCAACTGSAAGGACACACCTATGAAAGTGAAGTCAA	1921
QY	3146	LEUNARYCYELEUNGUGIYTYRTHMERASPHRAPHRIAPTRHETHRCYSGIUNYIS	3165
DB	1922	CTCAGANTCTCGAAGGTTATACATGATGATACAAATACAGTATACATTCCTTCACAAA	1981
QY	3166	AEPGIYATGTTPHPEPROGIUNARGLIESERCYSESPROLYSLYSCYSPROLEUPROGIUN	3185
DB	1982	GATGGTCCCTGGTCCCTCGAAGAAATCTCTCGCAAGTCTTAAAAAATGTCTCTCCGGAA	2041
QY	3186	ANILETHRHISILEUVALHISGIYASPAAPHESERVALAANARGIUNALSERVAL	3205
DB	2042	AACATTAACATATATACCTGTCAACAGGGACGATTTCAGTGTGAATAGCAAGTTCTGTG	2101
QY	3206	SERCYSALAGIUGIYTYRTHRPHEGIUGIYVALANILIESERVALCYSGIUNLEUAPGLY	3225
DB	2102	TCATGTGCAAGAGGGATATACCTTTGAGGAGATTAACATACAGTATCAGCTTGATGAGA	2161
QY	3226	THTRPGIUNPROPHESERASPOUSERCYSESPROVALSERCYSGIYLYSPROGIUN	3245
DB	2162	ACCTGGAGCAACCATTTCTCCGAGAAATCTTCAGTCCAGATTTCTTGTGGAAACCTGAA	2221
QY	3246	SERPROGIUNHISGIYPHEVALVALGISERLYSRYRTHRPHEGIUSERTHRIELIESTYR	3265
DB	2222	AGTCAGAAACATGATTTGTGGTGGCGACGTAAATACACTTTGAAGACACATATTTAT	2281
QY	3266	GLNCYSGIUNPROGIYTYRGILULENGUGIYANARGIUNAGVALCYSGINGIUNANARG	3285
DB	2282	CAGTGTGAGCTGTGGCTATGAACATGAAGGGGAAACGGGAAACGTGTCTCGAGGAACACA	2341
QY	3286	GLNTRPSERGILYGLVALAIIIECYBISYSGIUNTRARGCYSGIUNTRPROLEUGIUNPHE	3305
DB	2342	CAGTGAAGTGGAGGGGTGGCAATATGCAAAGACCAAGTGTGAACCTCCACTTGAATTT	2401
QY	3306	LEUNENGIYLYSALASPILIEGIUNANARGTHRTHRGIYPROAHNVALIYTRSERCYIS	3325
DB	2402	CTCAATGGGAAACGTCACATTTGAATAACAGACACTGGAACCAACGTGTATATTCCTGC	2461
QY	3326	ANANARGIYTYRSEULENGUGIYPROSERGIUNALHISCYSTHNGIUNENGIYTHTRTP	3345
DB	2462	AACCAAGGCTTCAAGCTTGAAGGGGCACTGACGACACTCACAGAAAATGGAACCTGG	2521
QY	3346	SERHISPROVALPROLEUCYLYSPROAHNPROCYSPROVALPROPHEVALIIPROGIUN	3365
DB	2522	AACCAACCAAGTCCCTCTCTCGAACAACAAATCCATGCCCTGTTCCTTTGTGATTCGCCAG	2581
QY	3366	ANALALEULEUSERGIUNYSGIUNPHEITYVALAEPGINANVALSERILYSCYSARG	3385
DB	2582	AATGCTCTGCTGTCTGAAAAGGAGTTTATGTATGTATGATCAAGAAATGTCCATCAAAATGAGG	2641
QY	3386	GIUNYIYPHEULEUNGILCYHISGLYILEIETHRCYAHNPROAEPGIUNTRTPHTR	3405

Db 1561 GAGGACATTGACCTGACAGAGAGAACTGACCAAAAAATATGCTGATATAT 1620
 Qy TyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAsnArgLeu 877
 Db 1621 TAGACATATGAAATGCTTTGCAATTGACAGAGTGGCTGGGGTGCAGCTAAATAGGCTG 1680
 Qy AspTyrSerTyrAspAspPheLeuAspThrValGlnGlnTyrAlaIleThrSerIleGlyValAsn 897
 Db 1681 GATTACTCTTACAGATGACTCTCTGACACCTGACAAAGAAACCAAGACCAAGATGGCAAT 1740
 Qy AlaIysSerSerArgIleLeuArgSerAlaProLeuSerAspTyrTyrIleLeuLeuIle 917
 Db 1741 GCCAGGCTCTCACGGATTAAAGAAAGTCCCATTTATCTGACTTAAATTAAGTTATTT 1800
 Qy PheAsnIleThrAlaSerValProLeuProAspGluArgAsnArgThrLeuGlnTyrGlu 937
 Db 1801 TTTAACTATCAGCTAGTGTCCCATTAACCGATGAAGAAATGATACCTTGAAGTGA 1860
 Qy AsnGlnGlnArgLeuGlnThrLeuGlnTyrIleThrAsnIleLeuValArgThrLeu 957
 Db 1861 AATCAGCAACGACTCTTCACACATTTGAACTATCAGAAATTAATGAAAGACTCTC 1920
 Qy AsnIysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsn 977
 Db 1921 AACAAAGACCCCATGTATTTCTTTCACTGATCAGAAATTAATGATGCGACAGCAAT 1980
 Qy SerLeuGlnTyrIleValAlaSerProPheCysArgProGlySerValLeuArgGlyVal 997
 Db 1981 TCATTAGAAACAAAAAGGCTTCCCTTCGCAACAGGCTCAGTGTGAGAGGGGCT 2040
 Qy MetCysValAsnGlyProLeuGlyThrTyrTyrAsnLeuGlnIlePheThrCysGlnSer 1017
 Db 2041 ATGATGTCATATGCTCTTGGAACTATTAATTCGAAACATTTCACTGTAAGAC 2100
 Qy CysArgIleGlySerTyrGlnAspGluGlnGlnLeuGlnCysLeuGlnCysProSer 1037
 Db 2101 TCCCGATCGGATCTTATCAAGTGAAGAGGCACTTGAATGCAAGCTTTGGCCCTCT 2160
 Qy GlyMetTyrThrGlnTyrIleIleSerArgAsnIleSerAspCysValAlaGlnCysIle 1057
 Db 2161 GGGATGTCACGGAATATATTCATTCAGAAACATCTCGATTTGAAGCTCAGGTAA 2220
 Qy GlnGlnTyrTyrSerTyrSerGlyLeuGlnTyrCysGlnSerCysProLeuGlnTyrTyr 1077
 Db 2221 CAAAGCACTTACTCAGTACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2280
 Qy GlnProIysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValIys 1097
 Db 2281 CAGCCAAATTTGTTCCCGAGCTGCTCTGTCGCAAGAAACACTCAACTGAGAA 2340
 Qy ArgGlyAlaValAlaAsnIleSerAlaCysGlyValProCysProGlnGlnIlyPheSerArg 1117
 Db 2341 AGAGAGCCGTTGACATTTCTGCAATGAGTTCCTTCGCAAGAAAGTTCGCT 2400
 Qy SerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnIleGlyIys 1137
 Db 2401 TCTGGTTAAATGCTCTGCTACCAATCTCTGAGCTATTCACACTTAATGAGGAAG 2460
 Qy AlaPheCysLeuAlaCysProPheTyrGlyTyrThrProPheAlaGlySerArgSerIle 1157
 Db 2461 GCCTTCTGCTGCTGCTGCTCTTTATGAACTACCCCATTTGCTGCTGCTGCTGCTGCT 2520
 Qy ThrGlnCysSerSerPheSerSerThrPheSerAlaAlaGlnGlnIleValProPro 1177
 Db 2521 ACAGATGTTCAAGTTTGTGTTCACTTCTCAGCGCAGAGGAAAGTGTGGCCCT 2580
 Qy AlaSerLeuGlyHisIleLeuValArgHisGlnIleSerSerGlnValPheHisGlnCys 1197
 Db 2581 GCTCTCTTGGACATATTAAGAGGCAATGAAATCAGAGTCAAGTTTCCATGAAATG 2640
 Qy PhePheAsnProCysHisAsnSerGlyTyrCysGlnGlnIleGlnIlyArgGlyTyrValCys 1217
 Db 2641 TTCTTTAACCTTGGCAGATATGTGAACCTGCAACCACTTGGGCTGTATGTTGT 2700

Qy 1218 LeuCysProLeuGlnTyrTyrThrGlyLeuIlyPheCysGlnTyrAspIleAspGluCysSerPro 1237
 Db 2701 CTCTGTCACCTTGATATACAGGCTTAAAGTGTAAACAGATCAATGAGTGCAGCCCA 2760
 Qy 1238 LeuProCysLeuAsnAsnGlnValCysIlyAspLeuValGlyGlnPheIleCysGlnCys 1257
 Db 2761 CTGCTTGCCTCAACCAATGAGATTGTGTAAAGACCTAGTGTGGGAAATCATTTGTATGTC 2820
 Qy 1258 ProSerGlyTyrTyrThrGlyGlnArgCysGlnGlnIleAsnGlnCysSerSerPro 1277
 Db 2821 CCATCAGGTTACAGAGTCAAGGTGTAAGAAATATTAATGAGTGAAGTCTCAAGTCT 2880
 Qy 1278 CysLeuAsnIlyGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValIys 1297
 Db 2881 TGTTTAATTAAGAAATCTGTCTTCAATGCTGTGTGCTGCTATCGTGCACATGTGAAA 2940
 Qy 1298 GlyPheValGly 1301
 Db 2941 GAATTTGTAGGT 2952
 Db
 RESULT 23
 AAS0387
 ID AAS03887 standard; cDNA, 2929 BP.
 AC AAS03887;
 XX 29-AUG-2001 (first entry)
 DT
 XX
 DB Human secreted protein gene #6.
 XX
 KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;
 KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;
 KW nervous system disorder; bacterial infection; viral infection; 89;
 KW fungal infection; ocular disorder; wound healing; tissue regeneration;
 KW epithelial cell proliferation; skin aging; chemotaxis; IgG Fc region.
 XX
 OS Homo sapiens.
 XX
 PN W0200123598-A1.
 XX
 PD 05-APR-2001.
 XX
 PF 26-SBP-2000; 2000WO-US026324.
 XX
 PR 27-SBP-1999; 99US-0155807P.
 XX
 PA (HDMA-) HDMA GENOME SCI INC.
 XX
 PI Komatsoulis G, Ruben SM, Rosen CA;
 XX
 DR WPI: 2001-281684/29.
 DR P-PSDB; AAU01931.
 XX
 PT Forty nucleic acid molecules encoding human secreted proteins, useful in
 PT the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases.
 XX
 PS Disclosure; Page 447-448; 518pp; English.
 XX
 CC Sequences AAS03873-AAS03922 represent isolated nucleic acid molecules and
 CC PCR primers of the invention. acid of the invention. Secreted proteins
 CC and their related nucleic acids can be used in the diagnosis of or
 CC susceptibility to a pathological condition by determining the presence or
 CC absence of a mutation in a nucleic acid or the presence or amount of
 CC expression of a secreted protein. The sequences are used to prevent,
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. The antibodies to the
 CC polypeptides can also be used in alleviating symptoms associated with
 CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme
 CC linked immunosorbent assays (ELISA). The disorders include autoimmune
 CC diseases e.g. Rheumatoid arthritis, hyperproliferative disorders e.g.
 CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac

QY 1711 ThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerPheMetPheCys 1730
 DB |||||
 QY 1221 ACAGTAACCTTACCAAGTGCACCAATGCGTACTATCTATGGGTGACTCAAGAGATGTTCTGT 180
 DB |||||
 QY 1731 ThrAspAsnGlySerTTPAsnGlyValSerProSerCysLeuAspValAspGluCysAla 1750
 DB |||||
 DB 181 ACAGATATATGGAGCTGGAAAGCGGTTTCACCAATCCGCTTGAATGTCAGTAAGTGTGCA 240
 QY 1751 ValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGlySerTyr1LeCys 1770
 DB 241 GTTGGATCAGATATTTGATGAGCAGTCTTCTGCTGAACGTAGATGATCTTACATATGT 300
 QY 1771 SerCysValProProTyrThrGlyAspGlyValAsnGlyValAspGluPro1LeCysCysAla 1790
 DB 301 TCATGTGTCCACCGTACACAGAGAGATGGGAAAACTGTGCAGAACTATTAATGATAG 360
 QY 1791 AlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu1LeTyrThrValGlyAla 1810
 DB 361 GCTCCAGAAATCCGAAAAATGCGCACTCTCAGGTAGATTAATACAGTGTGCGGAA 420
 QY 1811 ValThrPheSerCysGlnGlnGlyTyrGlnLeuMetGlyValThrTyrValLeThrCysLeu 1830
 DB 421 GTCAACATTTTCGTCTCAGAAAGATACCAAGTTGATGGAGATACCAAAATCACATGTTTG 480
 QY 1831 GluSerGlyGluTTPAsnHisLeu1LeProTyrCysAlaValAspSerGlyValPro 1850
 DB 481 GAGCTCGAGAAATGGAATCATCTAATATCATATGTAAAGCTGTATTCAATGTGTAAACCG 540
 QY 1851 Ala1LeProGluAsnGlyCys1LeGluGluLeuAlaPheThrPheGlySerTyrValThr 1870
 DB 541 GCTATTCCAAAAATGGTGTGATTTAGAGATTAACATTTACTTTTGGAGAAAGTGACA 600
 QY 1871 TyrArgCysAsnGlyGlyTyrThrLeuAlaGlyAspGlyGlySerSerCysLeuAlaAsn 1890
 DB 601 TATAGGATTAATAAGATATACTCTGCGCGGTATTAAGAAATCATCTGCTTGTCTGATAC 660
 QY 1891 SerSerTTPSerHisSerProProValCysGluProVal1LeCysSerSerProGluAsn 1910
 DB 661 AGTTCTGTGAGTATTCCTCCCTGCTGTGTGAAACCACTGAAAGTGTCTAGTCCGAAAT 720
 QY 1911 IleAsnAsnGlyTyrTyr1LeuSerGlyLeuThrTyrLeuSerThrAlaSerTyrSer 1930
 DB 721 ATAAATATATGAAATATATTTTGAATGGCTTACCTTACTTACTGATCATATATCA 780
 QY 1931 CysAspThrGlyTyrSerLeuGlnGlyProSer1Le1LeGluCysThrAlaSerGly1Le 1950
 DB 781 TCGGATACAGGATACAGCTTACAGGCGCTTCCATATATGAAATGACAGCGCTTGGCATC 840
 QY 1951 TrpAspArgAlaProProAlaCysHisLeuValPheCysGlyGluProProAla1Le1Le 1970
 DB 841 TGGGACAGACCGGCACTGCTGTGTCTGTCTGTGTGTGAAACCACTGCGCATAA 900
 QY 1971 AspAlaVal1LeThrGlyAsnAsnPheThrPheArgAsnThrVal1ThrTyrThrCysAla 1990
 DB 901 GATGCTGTCAATTCAGGGAATTAACCTTACCTTACGAAACACCGTCACTTACCTTGCAA 960
 QY 1991 GluGlyTyrThrLeuAlaGlyLeuAspThr1LeGluCysLeuAlaAspGlyValSer 2010
 DB 961 GAAGGCTATATCTGTGCTGTGCTTTCACACCATTAATGCGCGCAAGGAAATGGAGT 1020
 QY 2011 ArgSerAspGlnGlnCysLeuAlaValSerCysAspGluProPro1LeValAspHisAla 2030
 DB 1021 AGAAGTACACGACAGTCCGTGGCTGTCTCTGTATAGCACTTGTGTGAAACCAAGCC 1080
 QY 2031 SerProGluThrAlaHisArgLeuPheGlyAsp1LeAlaPheTyrTyrCysSerAspGly 2050
 DB 1081 TCTCCAGAGCTGCCATGCGCTTGTGAGACATTTGCAATTTCTACTGCTGTGAATGTT 1140
 QY 2051 TyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyValSerProPro 2070
 DB 1141 TACAGCTTACAGACATTTCCAGCTTCTGTGCAATGCCAGCGGCAAGTGTGTATCCCCCA 1200

QY 2071 GluGluGlnAspMetProArgCysAlaAlaHisPheCysGlyValProProSerValSer 2090
 DB |||||
 DB 1201 GAAGGTCAAGACAGCCCCGTTGATAGCTATTTGTGAAAACTTCCATCGGATTC 1260
 QY 2091 TyrSer1LeuGlnSerValSer1LeAlaLysPheAlaAlaGlySerValValSerPhe 2110
 DB 1261 TATAGCATCTTGGATCTGTGAGCAAGCAAAATTTGCAAGTGGCTGATGTCATTT 1320
 QY 2111 LysCysMetGluGlyPheValLeuAsnThrSerAlaLys1LeGluCysMetArgGlyGly 2130
 DB 1321 AAATGCATGAAAGCTTTGTACTAGAACCTCAGCAAAATGTAATGATATGAGAGTGG 1380
 QY 2131 GlnTTPAsnProSerProMetSer1LeGlnCys1LeProValArgCysGlyGluProPro 2150
 DB 1381 CAGTGAACCTTCTCCCATGTTCATCCAGTGCATCTGTGCGGTGTGAGAGCCACCA 1440
 QY 2151 Ser1LeMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyr 2170
 DB 1441 AGCATCATGATGGCTATGCAAGTGGATCAAACTACAGTTTGGAGCATGTGGCTTAC 1500
 QY 2171 SerCysAsnGlyGlyPheTyr1LeGlyGlyGluLysSerThrCysGluAlaThrGly 2190
 DB 1501 AGTGCAAACAGGCGTTCTATCATCAAAAGGGAAGACACCTGCGAAGCCACAGGG 1560
 QY 2191 GlnTTPSerSerPro1LeProThrCysHisProValSerCysGlyGluProProVal 2210
 DB 1561 CAGTGAAGTATCTTATACCAAGTGCACACCGGATTTGTGTGAAACCACTTAAGTT 1620
 QY 2211 GluAsnGlyPheLeuGlnHisThrThrGlyArg1LePheGluSerGluValArgTyrGln 2230
 DB 1621 GAGATGGCTTTCTGGAGCATACAACTGGCAGATCTTTGAGAGTGAAGTGAAGTATCAG 1680
 QY 2231 CysAsnProGlyTyrTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHis 2250
 DB 1681 TGTAAACCGGCTTAAAGTCAAGTGCAGAAATCTGATTTGTCTGCGCAAGCCAAATGCGCAC 1740
 QY 2251 TrpHisSerGlySerProLeuMetCysValProLeuAspCysGlyValProProPro1Le 2270
 DB 1741 TGGCAAGTGAATCCCTGTATGTGTGTCTCTGACGTGAAACCTCCCGCATC 1800
 QY 2271 GlnAsnGlyPheMetGlyGlyGluAsnPheGluValGlySerLysValGlnPhePheCys 2290
 DB 1801 CAGATGGCTTCAAGAAAGGAAAGAACTTTGAAAGAGGCTCAAGGTTCAAGTTTCTGT 1860
 QY 2291 AsnGluGlyTyrGluLeuValGlyAspSerSerTTPThrCysGlnLysSerGlyValSer 2310
 DB 1861 AATGAGGTTATGAGCTGTGTGTGACAGTTCTTGACATGTCAAGAAATCGGCAAAATGG 1920
 QY 2311 AsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuGln 2330
 DB 1921 AATAAGAAATCAAAATCCAAAGTGCATGCGCAAGTCCCAAGAGCCGCGCTTGGAA 1980
 QY 2331 AsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLys 2350
 DB 1981 AACCAAGTATTAAGAGATTTGACCAAGGATTAAGAGTGTGACATTTTCCGTAAA 2040
 QY 2351 GluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTTPAsn 2370
 DB 2041 GAAGGAGATGTCCGCAAGGCGCTGTGTCTGTGAAATGCTTGCAATCCACAGATGGAAT 2100
 QY 2371 AsnSerPheProValCysLysAlaValLeuCysThrProProProLeu1LeSerPheGly 2390
 DB 2101 GACTCTTCCCTGTGTGTAATGTGTCTTTGACCCCACTCCCTTAATTTCTTGTGT 2160
 QY 2391 ValPro1LeProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGly 2410
 DB 2161 GTCCCATTTCTTCTGTGCTCTTCAATTTGGAAGTACGTCAAGATATCTTGTGTAGGT 2220
 QY 2411 GlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTTPSerSer 2430
 DB 2221 GGGTTTTTCCTAAGAGAAATTTTACCACTCTGCGCAACTGTATGGACCTGAGGCTCT 2280
 QY 2431 ProLeuProGluCysValProValGluCysProGlnProGluGlu1LeProAsnGly1Le 2450

DB 2281 CCACGCGAGATGTTCCATGATGTCCTCCACCTGAGGAAATCCCAATGATC 2340
 QY 2451 ILeaPValGInGlyVLeuAlaThyLeuSerThrAlaLeuThyThyCysLysProGlyPhe 2470
 DB 2341 ATTGATGTCAGAGGCTTCCCTATCTCAGACAGAGCTCTCTATCTGCAAGCAGGCTTT 2400
 QY 2471 GLeuValGlyAsnThrThyLeuCysGlyValAsnGlyHisTrpLeuGlyGlyLys 2490
 DB 2401 GAATTGGTGGGAAATATCTACCACTCTTGTGAGGAAATATGTCATCGCTTGGAGGAAA 2460
 QY 2491 ProThrCysLysAlaAlaGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSer 2510
 DB 2461 CCACATGTAAAGCCATGAGTCCGAAACCCAGAGAAATTTGAATGCAAAATCTCT 2520
 QY 2511 TTTTAAAPLeuHisThyCysGlyValThyValThyThySerCysAsnArgLysPheArgLeu 2530
 DB 2521 TACACGAGACTACACTATGAGACAGACCGTTACTCTTGGCAACGAGGCTTTGGCTC 2580
 QY 2531 GInGlyProSerAlaLeuThrCysLeuGluThyGlyAspTrpAspValAsp-AlaProSe 2550
 DB 2581 GAAAGTCCAGTGCCTTGACCTGTATGAGACAGGCTGATGGATGTAGATTGCCCATC 2640
 QY 2550 TCYAsnAlaAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGlu 2568
 DB 2641 TTGCAATGACATCCACTGTGATTCCCAACACCATTTGA--AATGTTTGTAAAG 2693
 RESULT 25
 ABS51807
 ID ABS51807 standard; cDNA; 3262 BP.
 AC ABS51807;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 XX Human mdtc cDNA Incyte ID No: LI:202943.4:2001JUN12.
 DE
 XX Human; molecule for disease detection and treatment; MDDT; cancer;
 KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
 KW autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
 KW hepatotropic; immunosuppressive; antistimatic; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 XX MO200255738-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002MO-US001008.
 XX
 XX 12-JAN-2001; 2001US-0261622P.
 PR 16-JAN-2001; 2001US-0261865P.
 PR 17-JAN-2001; 2001US-0262208P.
 PR 17-JAN-2001; 2001US-0262209P.
 PR 17-JAN-2001; 2001US-026236P.
 PR 19-JAN-2001; 2001US-0263063P.
 PR 19-JAN-2001; 2001US-0263065P.
 PR 19-JAN-2001; 2001US-0263329P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 PI Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JT, Jones AL,
 PI Dam TC, Liu TF, Harris B, Flores V, Dafo A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstein BH, Peralta CH, David ME, Lewis SA;
 XX
 DR WPI: 2002-590679/63.
 DR P-PBDB: ABS70334.
 XX
 PT New disease detection and treatment molecule (MDDT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDT expression, e.g. autoimmune or inflammatory
 PT disorders.

XX
 PS Claim 1: Page 102-103; 129pp; English.
 XX
 CC The present invention relates to the isolation of novel human molecules
 CC for disease detection and treatment (MDDT), and the polynucleotide
 CC sequences (mdt) encoding them. The MDDT polypeptides may be used to
 CC screen for molecules that bind to, or are bound by the encoded
 CC polypeptides, and to develop a transcript image of a tissue or cell type.
 CC Probes comprising at least 20 nucleotides of the mdt polynucleotide may
 CC be used to assess the toxicity of a test compound. The MDDT polypeptides
 CC and mdt polynucleotides are useful in the diagnosis, study, prevention
 CC and treatment of diseases associated with the expression of molecules for
 CC disease detection and treatment. Such disorders include cell
 CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, or cancer),
 CC and autoimmune/inflammatory disorders (e.g. asthma, Crohn's disease, or
 CC multiple sclerosis). The mdt polynucleotides may also be used as
 CC molecule markers, in microarrays, and in somatic or germline gene
 CC therapy. ABS51779-ABS51814 encode the MDDT proteins of the invention
 XX
 SQ Sequence 3262 BP; 863 A; 760 C; 791 G; 848 T; 0 U; 0 Other:
 Alignment Scores:
 Pred. No.: 2, 03e-238 Length: 3262
 Score: 4930.00 Matches: 947
 Percent Similarity: 95.33% Conservative: 12
 Best Local Similarity: 94.14% Mismatches: 20
 Query Match: 24,68% Indels: 28
 DB: Gaps: 4
 US-09-977-053-4 (1-3571) x ABS51807 (1-3262)
 QY 1590 LysSerLeuAlaThySerCysProGluGluLeuSerLysGlyValLeuAlaTrpPro 1609
 DB 289 GAGTCACTGCTAACCCTCCCTGCGCCAGAGAACTCAGTAAAGAAACGTTGAGATGCTCT 348
 QY 1610 AspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysSerIlePheCysSer 1629
 DB 349 GATTCTTGTCAAGAAATGTGGGAAAGTGAATGATTCATTAGAGCAATTTTGTCT 408
 QY 1630 AspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLys 1649
 DB 409 GGTGCCCAACGCTTGAAGAGGTCAGTGCCTCATCTGAACTCATCTGAAGATTTAAAG 468
 QY 1650 ProGlySerLysValAsnLeuPheCysAspProGlyPheGluLeuValGlyAsnProVal 1669
 DB 469 CCAAGTTCACAAAGTCAATCTGTTCTGTGATCCAGGCTTCCAGCTGGTCGGAAACCTGTG 528
 QY 1670 GlnThrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgIleSer 1689
 DB 529 CAGTACTGTCTGAATTAAGACAGTGCAGACACCACTTCTCATCTGTGAACGATTAGC 588
 QY 1690 CysGlyValProProProLeuGluGluAsnGlyPheHisSerAlaAspAspPheTrpAlaGly 1709
 DB 589 TGTGGGTGTCACCTCTTTGAGAAATGAGCTTCCATTACAGCCATGCTCTATGCTGGC 648
 QY 1710 SerThrValThyTrpGlnCysAsnAsnGlyTrpTrpLeuGlnGlyAspSerArgMetPhe 1729
 DB 649 AGCACAGTAAACTACACAGTGCACCAATGCTACTATCTATGTGGGTACTCAGGAGATGTC 708
 QY 1730 CysThrAspAsnGlySer-TrpAsnGlyValSerProSerCysLeuAspValAspGluCys 1749
 DB 709 TGTACGATTAATGGAGCTAGAGAACGCGCTTTCACATCTCCGCGATGTGATAGTG 768
 QY 1749 AlaValGlySerAspCysSerGluHisAlaSerCysLeu-AsnValAspGlySerTrp 1769
 DB 769 TGCAGTTGATCAGATTGATGATGACATCTTCTCCCTGAAACGTAGATGATCTTACCA 828
 QY 1769 LeCysSerCysValAProProTrpThrGlnAspGlyLysAsnGlyAlaGluProIleLysC 1789
 DB 829 TATGTCATGTGTGCCACGTTACACAGAGATGGGAAATCTGTGCAAACTTAAAT 888
 QY 1789 Ys-LysAlaProGly--AsnProGluAsnGlyHisSerSerGlyGluIleTrpThr-Va 1807

Dh	889	GCTAAGGCTCCAGGCCAGAAATCCGGAAAAATGGCCACTCTCTCAGAGGAGATTATTAACAGT	948
Qy	1807	1GIyAla---AlaValThrPheSerCyseGlnGluGlyTyr-GlnLewMetGlyValThrL	1826
Dh	949	AGGTGCCCGAAGTGCATTATATTAGCTGTCAAGAAAGATATCCAGTATGATGGAGATACCA	1008
Qy	1826	YellEthrCyseLenglu-SerGlyGluTrpPheHisLeuLlePro-TyrCyblyValaVa	1845
Dh	1009	AAATCACTAGTTTGGAGTACTGGAGAAATGGAAATCATCTTAATACCAATTTGTAAAGCTGT	1068
Qy	1845	1SerCyseGlyLysPro-AlaIleProGluAnGlyCybLleGlu-GlnLewAla---Phe	1863
Dh	1069	TTCAATGTGTAAACCGACTATCTCAGAAAAATGGTTGATATGACGAGACTTACCACTTTT	1128
Qy	1864	ThrPhe---GlySerLysValThrTyrArgCyseAnLysGlyTyrThLewAlaGlyLysP	1882
Dh	1129	ACCTATTGGGCGCAGAAATGTACATATAGGTGTATATAAGATATATCTTGGCCGGTAT	1188
Qy	1883	LysGluSerSerCybLewAlaAnsSerTrpSerHisSerProValCyseGluPro	1902
Dh	1189	AAAGATCATCTCTGCTTGGCTAACAGTTCTTGGAGTCAATCCCTCCCTGTGTGTGAACA	1248
Qy	1903	ValLysCyseSerSerProGluAnsLleHisAnGlyLys-TyrLleLewSerGlyLthrTh	1922
Dh	1249	GTTAAGTGTCTAGTCCGACATATTACTATATGACACTATATATAGAGTGGGCTTAC	1308
Qy	1922	TrpLewSerThrAlaSer-TyrSerCyseAspThrGlyTyrSerLengluGlyProSer-	1941
Dh	1309	CTACCTTTCTACTGCATCATATATTCATGCAATACAGATATACAGCTTACAGGGCCCTTCC	1368
Qy	1942	LleIleGluCyethrAlaSerGlyLleTrpAspArgAlaPro-ProAlaCybHisLewVa	1961
Dh	1369	ATTATTTGAATGACACGGCTTCTGGGCATCTGGGACACAGGGGACCCCTGCTGCATCTGT	1428
Qy	1961	LpheYseGlyGluProProAlaIleLysAspAlaValLleThrGlyAsnAnpEthrPh	1981
Dh	1429	CTTCTGTGAGAACACCTGCATTAAGATATCTGTCAATACGGGAAATTAACCTTACTTT	1488
Qy	1981	eArgAnthrValThrTyrThrCybLysGluGlyTyrThrLewAlaGlyLewAspThrLl	2001
Dh	1489	CAGGACACCGTCACTTACCTTGCAGAAAGGCTATATCTGTGTGCTTGGACACAT	1548
Qy	2001	eGluCybLewAlaAspGlyLysTrpSerArgSerAspGlnGluCybLewAlaValSerCy	2021
Dh	1549	TGAATGCTGGGCGACGGCAAGTGAATGAGATGACACAGCATGCTCGGCTGTCTCTG	1608
Qy	2021	AspGluProProIleValAspHisAlaSerProGluThrAlaHisAlaGluPheGlyAs	2041
Dh	1609	TGATATGACACCCATTTGTGTGACACAGGCTTCTCCAGAGATCGCCATCGGCTTTCGAGA	1668
Qy	2041	PileAlaPheTyrTyrCyseSerAspGlyTyrSerLewAlaAspAnsSerGlnLewCy	2061
Dh	1669	CATTGCATTTCTACTCTCTCTGATGATTACAGCTTAGAGACATATCCAGCTTCTCTG	1728
Qy	2061	AsnAlaGlnGlyLysTrp-ValProProGluGlyGlnAspMetProArgCybLleAlaH	2081
Dh	1729	CAATGCCAGGGCAAGTGGGGATACCCCAAGAGATCAAGACATGCCCGCTGTATAGCTC	1788
Qy	2081	LepheCyseGlu-LysProProSerValSerTyrSerLleLengLysEthrValSerLysAla	2100
Dh	1789	ATTCTGTGAAAACCACTTCATCGGTTTCTATATACACTTTGGATCTGTGTAGCAAAACA	1848
Qy	2101	LysPheAlaAlaGlySerValValSerPheLysCyseMetGluGlyPheValLewAnsThr	2120
Dh	1849	AAATTTGACGCTGCTCACTGTGTGAGCTTATATATGATGAAAGCTTTGTACTGAAACC	1908
Qy	2121	SerAlaLysIleGluCyseMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln	2140
Dh	1909	TCAGCAAGATTTGAATGTATATGAGGTGGGAGTGAACCTTCCCATGTGCATTCAG	1968
Qy	2141	CybIleProValArgCyseGlyGluProProSerLleLewAnsGlyTyrAlaSerGlySer	2160
Dh	1969	TGCATATCCCTGTGCGGTGTGTGAGAGCACCAAGCATCATGAATGGCTATGCAAGTGAACA	2028

QY	2161	ASNTYISerPheGluValAmelValAlaTyrSerCyAsnIlysgIyPheTyrTlleYsgIy	2180
Db	2029	AACTACAGTYYTGAGGCCATGGTGGCTTACAGCTGCACACAGGGGTCTCACTCAAGAGG	2088
QY	2181	GIuIyIysIysSerThrCyAsnIuAlaIsthGIuIIntPsrSerProIlePsrOthrCyHis	2200
Db	2089	GAAGAAGAGCACCTCGAAGCCACAGGGCAGTGGAGTACTTATACGACCTGCCAC	2148
QY	2201	ProValSerCySglYgluProProIyValGIuAsnGIyPheIeuGIuHisThrThGIy	2220
Db	2149	COGGATCTTGTTGGTGAACCACTTAAGGTGAGAAATGGCTTTCGAGGACATCAACTGC	2208
QY	2221	ArgIlePheGluSerGIuValArgTyrGlnIuCyAsnProGIyTyrIysSerValIsgIy	2240
Db	2209	AGGATCTTGAAGATGAAGTTGAGTATCAGTGTACCCGGGCTATATAGTCAGTCGGA	2268
QY	2240	rProValPheValCySgluAlaAsnArgHisTrrPHisIserGIuSerProIleuNetCyAs	2260
Db	2269	TCTGTATTTTGTCTGCCAAGCCATGCCACTGGCACAAGTGAATCCCTCTGATGTGT	2328
QY	2260	lProIleuAspCySglYtyrProProIleGlnAsnGIyPheMetIySglYgluAsnPh	2280
Db	2329	TCTCTTCACATGTGAAACCTCCCCGATCCAAATGGCTTCATGAAAGAGAAAATT	2368
QY	2280	egIuValGIySerIyValGlnPhePheCyAsnGIuGIYrGIuIeuValGIyAsp	2300
Db	2389	TGAAGTAAAGGTCGAAAGTTCAGTTCCTGTATAGAGGTTATAGCTTGTGGTGCAG	2448
QY	2300	rSerTrrPThrCyAsnIyysSerGIyIystrPAsnIyIysSerAsnProIyCySmetPr	2320
Db	2449	TTCTTGACATAGTCAGAAATCTGGCAATATGAATTAAGATCAAAATGATGCAGTCC	2508
QY	2320	oAlIyCySerProIuProProIleuGIuAsnGIuIeuValIleuIySglIueuThr	2340
Db	2509	TGCCAAGGCCCAAGGCCGCCCTCTTGAAACCACTAGTATTAAGAAGTTGACAC	2568
QY	2340	rGIuValGIyValValThrPheSerCyIySglIuGIyHisValIleuGIuIyProSerVa	2360
Db	2569	CGAGGTAGAGTTGTGACATTTTCTCTTAAGAAGGCATGTCCTGCAGAGCCCTCTGT	2628
QY	2360	lIleuIyCySleuProSerGlnGlnTrrPAsnAspSerPheProValCyIyIyIleValI	2380
Db	2629	CCTGAATATGCTTGCCATCCAGCAATATGATCTTTCCCTGGTGTGAAGATGTCT	2688
QY	2380	uCySThrProProIleuIleSerPheGIyValProIleProSerSerAlaIeuHisph	2400
Db	2689	TTGTATCCCACTTCCCTAATTTCTTGGTGTCCCATTCCTTCTTCTCTCATATT	2748
QY	2400	egIySerThrValIys-TyrSer-CysValGIyIyPhePheIeu-ArgGIyAsnSerTh	2419
Db	2749	TGGAAGTACTGTCAAGGTATTTCTGAATGTAGTGGTGTTTTCTCAACAGAAATCTAC	2808
QY	2419	rThrIleuCyGlnProAspGIyIyHTrP-serSerProIleuProGIuCySValIProValG	2439
Db	2809	CACCTCTGCACCACTATGTGCACCTGAAGGCTCTCCACTGACAGAAATGTGTCCATAG	2868
QY	2439	IuCySProGIuProGIuGIuIleProAsnGIyIleIleAspValGIuGIyIleuAlaTyrT	2459
Db	2869	AATGTCCCAACCTGAAGAAATCCCAATGGAATCATGTATGTGCAGAGGCTTGCTATC	2928
QY	2459	eusSerThrAlaIeuTrrThrCySleYrProGIyPheGIuIeuValGIyAsnThrThrTrrL	2479
Db	2929	TCACACACAGCTCTCTAATACCTGCAGACAGCTTTGATATGTGGGAAATACACACCC	2988
QY	2479	euCySglYgluAsnGIyHisTrrPleuGIyGIyIyIyProThrCyIySAlaIleGIuCyS-	2498
Db	2989	TTTTTGTGAGAAATATGTCACTGTGCTTGAAGAAACCAATGTATTAAGCATTAAGTGCC	3048
QY	2499	IeuIyAspProIySglIuIleIeuAsnGIyIyPheSerTrrThrAspIeuHisTrrGIyGln	2518
Db	3049	GTGAACCCCAAGAGATTTTGAATGGAAATTTCTTTACCGAGACTTACACTTATGACAG	3108

```
QY 2519 ThrValThrTyrSerCysAsn-ArgGlyPheArgLeuGluGlyProSerAlaLeuThrCy 2538
    |||||
Db 3109 ACCGTTACTTACTTGGCAACGAGGCTTTCGGCTCGAAGGTCCAGTGCCCTTGACTG 3168
    |||||
QY 2538 aLeuGluThrGlyAspTTPAspValAsp-AlaProSerCysAsnAlaIleHisCysAspS 2558
    |||||
Db 3169 TTAGAGACAGGTGANTGGATGTAGATTGCCCATCTTGCAATGACATCCACTGTGANT 3228
    |||||
QY 2558 ePProGlnProIleGluAsnGlyPheValGlu 2568
    |||||
Db 3229 CCCCAACACCATTTGA--AATGGTTTGTAAAG 3258
    |||||
```

Search completed: May 9, 2004, 16:14:42
Job time : 3756.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: May 9, 2004, 12:50:18 ; Search time 14189.5 Seconds
(without alignments)
5626.549 Million cell updates/sec

Title: US-09-977-053-6
Sequence: 1 MFRLLAFCCWGLALVSGMAT.....GVTKITGLSGGNHLLPYC 1842

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues
Total number of hits satisfying chosen parameters: 6340544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODBL=frame+ p2n.model -DBV=xlp
-O/csgn2.1/uspro.spool_p/us09977053/runat.06052004.075942.18188/app.query.fasta.1.5710
-DB=GenBml -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCTL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=100
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=25 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09977053 @CGN 1.1 27696 @runat.06052004.075942.18188 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEJOINT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DBV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: gb_da.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_ov.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_stg.*
28: em_un.*

29: em_vl.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htg_hum.*
40: em_htg_mus.*
41: em_htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10208	99.9	10878	6 AR435509	AR435509 Sequence
2	10208	99.9	10878	6 AX375161	AX375161 Sequence
3	9862	96.5	11152	6 AX686198	AX686198 Sequence
4	9812	96.0	11158	6 AX686196	AX686196 Sequence
5	8673.5	84.9	11289	10 AF206329	AF206329 Mus muscu
6	8652.5	84.7	11230	6 AR435510	AR435510 Sequence
7	8652.5	84.7	11230	6 AX375163	AX375163 Sequence
8	7537	73.8	5124	6 AX880905	AX880905 Sequence
9	7537	73.8	5124	6 BD158659	BD158659 Primer fo
10	7537	73.8	5124	9 AK027870	AK027870 Homo sapi
11	4776	46.7	3253	6 AK122605	AK122605 Homo sapi
12	4776	46.7	3253	9 AX880735	AX880735 Sequence
13	3073	30.1	1969	6 BD158559	BD158559 Primer fo
14	3073	30.1	1969	6 AK023591	AK023591 Homo sapi
15	2374	23.2	1760	6 BD229966	BD229966 Method an
16	1447	14.2	765	6 AX870332	AX870332 Sequence
17	1447	14.2	765	6 BD150394	BD150394 Primer fo
18	1341	13.1	706	6 AX677660	AX677660 Sequence
19	1328	13.0	18583	2 AC018297	AC018297 Drosophi
20	1328	13.0	148780	3 AC008326	AC008326 Drosophi
21	1328	13.0	174287	3 AC007977	AC007977 Drosophi
22	1328	13.0	270766	3 AE003615	AE003615 Drosophi
23	1304	12.8	801	6 AX869682	AX869682 Sequence
24	1304	12.8	801	6 BD149744	BD149744 Primer fo
25	1237	12.1	3262	6 BX548016	BX548016 Danio rer
26	1004	9.8	167023	2 BD055344	BD055344 Secretd
27	998	9.8	598	6 AX080445	AX080445 Sequence
28	983	9.6	675	6 AF152001	AF152001 Danio rer
29	927	9.1	7931	5 AF152001	AF152001 Mus muscu
30	922	9.0	7579	10 D32210	D32210 Mus muscu
31	914.5	8.9	6728	6 AX528265	AX528265 Sequence
32	914.5	8.9	9722	9 AF308601	AF308601 Homo sapi
33	914.5	8.9	9723	6 AR102329	AR102329 Sequence
34	914.5	8.9	9723	6 AX357059	AX357059 Sequence
35	914.5	8.9	9723	6 AX417417	AX417417 Sequence
36	914.5	8.9	11189	9 AF315356	AF315356 Homo sapi
37	913.5	8.9	8287	10 R415024	R415024 Rattus norv
38	913	8.9	96779	9 AL592463	AL592463 Human DNA
39	903.5	8.8	9166	5 XELXOTCH	XELXOTCH
40	901	8.8	8503	3 LCUS58977	LCUS58977 Lucilia cup
41	893.5	8.7	7410	6 AX528263	AX528263 Sequence
42	888.5	8.7	8449	3 AF537369	AF537369 Boophilus
43	887.5	8.7	7885	5 AB095016	AB095016 Cyrops py
44	886.5	8.7	7471	5 BRNOTCH	BRNOTCH
45	886	8.7	9193	10 AF508809	AF508809 Mus muscu
46	877.5	8.6	11467	6 AX779966	AX779966 Sequence
47	872	8.5	7596	6 AX695382	AX695382 Sequence
48	872	8.5	8064	6 AX695381	AX695381 Sequence
49	872	8.5	8064	10 MNNOTCH	MNNOTCH
50	869.5	8.5	8221	10 RRNOTCH	RRNOTCH
					AX57405 R.rattus mr

51	869	8.5	7693	6	AX921793	AX921793 Sequence
52	868	8.5	7575	3	BPAMPHINO	Y12539 Brachyoslo
53	867	8.5	3362	3	SUSEGPI	L08692 Strangyloce
54	865.5	8.5	7646	3	AF000634	AF000634 Lytechinu
55	853	8.3	7332	6	AX394722	AX394722 Sequence
56	853	8.3	7332	9	HUNTANI	M73880 Human TAN-1
57	853	8.3	7693	9	AF308602	AF308602 Homo sapi
58	847	8.3	10452	3	DRONOT05	MI6152 Drosophila
59	847	8.3	17137	3	DRONOTCH03	K03508 D.melanogas
60	847	8.3	27874	2	AC014049	AC014049 Drosophila
61	847	8.3	40933	3	DMC140G11	AL035395 Drosophila
62	847	8.3	110000	3	AB003426_2	Continuation (3 of
63	847	8.3	179069	2	AC104144	AC104144 Drosophila
64	847	8.3	182387	3	AC116543	AC116543 Drosophila
65	845.5	8.2	7319	10	AF164486	AF164486 Rattus no
66	841.5	8.2	7943	10	MMNOT0C	X74760 M. musculus
67	838	8.2	763	6	AR220869	AR220869 Sequence
68	837.5	8.2	2460	6	AX351023	AX351023 Sequence
69	837.5	8.2	5516	3	AF239608	AF239608 Drosophila
70	837	8.2	8298	3	AB001327	AB001327 Halocynthia
71	831.5	8.1	7615	6	AX695384	AX695384 Sequence
72	829	8.1	8091	6	A69561	A69561 Sequence 1
73	829	8.1	8091	6	AR300200	AR300200 Sequence
74	829	8.1	8091	6	BD006120	BD006120 Gene invro
75	829	8.1	8091	6	HS097669	U97669 Homo sapien
76	829	8.1	8257	6	AR220824	AR220824 Sequence
77	817.5	8.0	2229	3	AF397902	AF397902 Podocoryn
78	816.5	8.0	3609	10	MMNOTB	X68279 M. musculus
79	782	7.7	2959	5	AB027453	AB027453 Xenopus 1
80	764.5	7.5	6677	6	AR199052	AR199052 Sequence
81	764.5	7.5	6677	6	AR207806	AR207806 Sequence
82	764.5	7.5	6677	10	MMU43691	U43691 Mus musculu
83	762.5	7.5	34425	3	HSU95299	U61946 Caenorhabdi
84	759	7.4	6122	9	HSU95299	U95299 Human Notch
85	758.5	7.4	1483	3	SUSGR	M17421 Strongyloce
86	758.5	7.4	6532	10	MUSINT3MAM	M60456 Mus musculu
87	751.5	7.4	5575	10	RATJAPR	L38483 Rattus norv
88	749	7.3	4208	6	AR169937	AR169937 Sequence
89	749	7.3	4208	6	AR119902	AR119902 Sequence
90	745	7.3	3974	6	AX565642	AX565642 Sequence
91	745	7.3	3974	6	AX597042	AX597042 Sequence
92	745	7.3	3974	6	AX701333	AX701333 Sequence
93	745	7.3	3974	10	AF171092	AF171092 Mus muscu
94	745	7.3	4855	6	AX375031	AX375031 Sequence
95	745	7.3	4855	9	HSU61276	U61276 Human trans
96	745	7.3	5377	10	BC058675	BC058675 Mus muscu
97	745	7.3	6464	6	AR033974	AR033974 Sequence
98	745	7.3	6464	6	AR095920	AR095920 Sequence
99	744	7.3	4208	6	AR182312	AR182312 Sequence
100	744	7.3	5445	9	HSU7720	U77720 Human trans

ALIGNMENTS

RESULT 1	AR435509	10878 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	Sequence 1 from patent US 6656707.				
DEFINITION	AR435509				
ACCESSION	AR435509				
VERSION	AR435509.1	GI:40198412			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 10878)				
AUTHORS	Welcher, A.A. and Elliott, G.S.				
TITLE	C3b/C4b complement receptor-like molecules and uses thereof				
JOURNAL	Patent: US 6656707-A 1 02-DEC-2003;				
FEATURES	Location/Qualifiers				
source	1..10878				
	/organism="unknown"				
	/mol_type="genomic DNA"				

ORIGIN

Alignment Scores:				Pred. No.:				Length:			
Score:				10208.00				Matches:			
Percent Similarity:				99.95%				Conservative:			
Best Local Similarity:				99.78%				Mismatches:			
Query Match:				99.89%				Indels:			
DB:				6				Gaps:			
US-09-977-053-6 (1-1842) x AR435509 (1-10878)											
QY	1	MetTTPProAArgLeuAlaPheCysCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr	20								
DB	11	ATGTGGCTCTCGCTGCTTTGTGTGCTGGGCTGGCCCTGCTTCGGGCTGGGAGCC	70								
QY	21	PheGlnGlnMetSerProSerArgAenPheSerPheArgLeuPheProGlnThrAlaPro	40								
DB	71	TTTCAGCAATATGCTCCCGTGGCAATTTCAAGTTTCGCTTCCCGAGACGGGCCC	130								
QY	41	GlyAlaProGlySerTleProAlaProAlaProGlyAspGlnAlaGlySerArg	60								
DB	131	GGGGCCCCGGAGATATCCCGCGCCCGCTCCGCGACAGACGGCGGAGACGA	190								
QY	61	ValGlnArgLeuGlyValPheArgArgArgValArgLeuLeuArgGlnLeuSerGln	80								
DB	191	GTGAGACGGCTGGCCAGGGCTTCCGCGACCGCTGCTGCGGAGCTCAGCAG	250								
QY	81	ArgLeuGlnLeuValPheLeuValAlaAspAspSerSerValGlyGlnValaAspArg	100								
DB	251	CGCTGGAGCTTCTCTTCGCTGGATGATTTCTTCAAGGTGGGCAAGTCACTTCGC	310								
QY	101	SerGlnLeuMetCysPheValArgGlyLeuLeuSerAspPheProValValProThrAlaThr	120								
DB	311	AGGACCTCATGTTCTCCGCAAGCTGCTCCGACTCCCTGGTGGCCAGCGCAGC	370								
QY	121	ArgValAlaAlaIleValPheSerSerArgAsnTyrValValProArgValaAspTyrIle	140								
DB	371	CGGTGGCCATCTGACCTTCTGTCAGAACTAGGTGTGCGCGGTGCAATATATC	430								
QY	141	SerThrArgArgAlaArgGlnIleValCysAlaLeuLeuGlnGlnIleProAlaIle	160								
DB	431	TTCACCCGCGCGCGCGGCGACGACAGTGGCGCTCTCTTCCAAGATCCCTGCATC	490								
QY	161	SerTyrArgGlyGlyGlyTyrTyrTyrGlyValPheGlnGlnAlaGlnIleLeu	180								
DB	491	TCTTACCGAGTGGCGGCACTTACACAGGCGGCTTCCAGAGCCGCGCAATCTT	550								
QY	181	LeuIleAlaArgGlnAsnSerThrArgValValPheLeuIleThrAspGlyTyrSerAsn	200								
DB	551	CTTCATGCTAGAGAAATCAACAAAGTTGATTTCTCATGACTGATGATATTCAT	610								
QY	201	GlyGlyAspProArgProIleAlaIleSerLeuArgAspSerGlyValGlnIlePheThr	220								
DB	611	GGGGAGACCTTGACCAATTCAGCGCTCATCGAGATTCAAGAGTGGAGATCTTCACT	670								
QY	221	PheGlyTleTTPGlnGlyAsnIleArgGlnLeuAsnAspMetAlaSerThrProGlyGln	240								
DB	671	TTTGGCATATGGCAAGGAACATTCCAGAGCTAAATGACATGCTTCCACCCCAAGAG	730								
QY	241	GlnIleCysTyrLeuLeuIleSerPheGlnGlnPheGlnIleLeuAlaArgAlaLeu	260								
DB	731	GAGACCTGTATCTGTACACAGTTTGAAGATTGAGGCTTTAGCTCCCGGACATTG	790								
QY	261	IleGlnAspLeuProSerGlySerPheIleGlnAspAspMetValIleCysSerTyrLeu	280								
DB	791	CATGAAGATATACCTTCTGGAGTTTATTCAGATGATATGCTCATCTCTTATCTT	850								
QY	281	CysAspGlnGlyGlyAspCysCysAspArgMetGlySerCysGlySerGlyThrIleThr	300								
DB	851	TGTGATGAAGGCAAGACCTGCTGTGACCGAATGGAGAGCTGCATAATGTGGACACACACA	910								
QY	301	GlyIlePheGlnGlyGlyIleCysGlyValTyrTyrTyrGlyValGlyGlnTyrGlnCys	320								

```
Db 911 GGCCATTGTGAGTCATCTGTGAAAAAGGGTATTACGGGAAGGTCGCACTATGAAATGC 970
Qy 321 ThrAlaCyProSerGlyThrTyLysProGluGlySerProGlyGlyIleSerSerCys 340
Db 971 ACAGCTTGGCCATCGGGGACATACAAACGTAAGGCTCACAGAGAAATCACAGATGTC 1030
Qy 341 IleProCyProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
Db 1031 ATTCCATGTCCTGATGAAATATCACACTCTCCACTGTGAAGCACTCCCTGAAAGACTGT 1090
Qy 361 ValCyAspGluGlyTyTrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
Db 1091 GTCTGAGAGAGAGATACAGGGCATCTGGCCAGACCTGTGAACCTTGCCACTGCCCTGCC 1150
Qy 381 LeuLysProProGluAsnGlyTyRPhelIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
Db 1151 CTGAAACCTCCCGAAAATGGTTACTTTATCCAAAACCTTGCAACCAACCTTCATGCA 1210
Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
Db 1211 GCCTGTGGGGTCCGATGTCACCTCGATTTGATCTGTGGAAAGCAGCATCATCTTATGT 1270
Qy 421 LeuProAsnGlyLeuLeuTrpSerGlySerGlySerTyCysArgValArgThrCysProHis 440
Db 1271 CTACCCAAATGGTTGTGTGTCGGTTCAGAGAGCTACTGCAAGATGAAGAAACATGCTCAT 1330
Qy 441 LeuArgGlnProLysHisIleGlyHisIleSerCysSerThrArgGluMetLeuTyLysThr 460
Db 1331 CTCGGCAGCCGAAACATGCGCACATCAGCTGTTCTACAAAGGAAAGTTATATACACA 1390
Qy 461 ThrCysLeuValAlaCysAspGluGlyTyTrArgLeuGlyIleSerAspLeuThrCys 480
Db 1391 ACAATGTTGGTTCCTGCTGATGAAAGGGTATACAGACTAAGAGCAGTGAATCACTTACTGT 1450
Qy 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGlnValGlyIleCysSerThr 500
Db 1451 CAAGGAACAGCCAGTGGAGGACCAAGAACCCGGGTGTGAGAGCCCACTGTTCACC 1510
Qy 501 PheGlnMetProLysAspValIleIleSerProHisAsnGlyLysGlnProAlaLys 520
Db 1511 TTTCAATGCCCAAAATATCATCATATCCCCCAACACTGTGGCAAGCCAGCCAAA 1570
Qy 521 PheGlyThrIleCysTyTrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGln 540
Db 1571 TTGGGACCATCTGCTATGTAAGTGGCCGCAAGGGTTCAATTATCTGGAAGTCAAAAGAA 1630
Qy 541 MetLeuAspCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
Db 1631 ATGCTAGAGATGACCACTCTGAAAAATGAAATGTCGGAAGTTCAGGCACTGTGTAA 1690
Qy 561 AspValGlnAlaProGlnIleAsnCysProLysAspIleGlnAlaLysThrLeuGln 580
Db 1691 GACGTGAGGCTCTCTCAATCACTGCTTCAAGACATAGAGGTAAAGACTGTGAACAG 1750
Qy 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyLys 600
Db 1751 CAAGATTCTGCAATGTTACTGTGAGATTCCAAAGCTAAAGCAACTCTGGTGAAG 1810
Qy 601 ValSerValHisValHisProAlaPheThrProProTyTrLeuPheProIleGlyAspVal 620
Db 1811 GTGTACGTCCACGTTCATCCAGCTTTCACCCCACTTACCTTTCCAGATGAGATGT 1870
Qy 621 AlaIleValTyTrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
Db 1871 GCATCTGATACACGGCAACTGACTATCGGCAACCAAGCCAGCTGCAATTTCCATATC 1930
Qy 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
Db 1931 AAGGTATATGATGACAAACCACTGTATGACTGTGGCAGATCTCCACTCCCGTCCAG 1990
Qy 661 ValSerGluLysValHisAlaIleSerTrpAspGluProGlnPheSerAspAsnSerGly 680
Db 1991 GTCTCGAAGAGATACATGCGCAAGCTGGAGTGAAGCTCATGTTCTCAGCAACTCAGGG 2050

Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyLeuThr 700
Db 2051 GCTGAATTGTCTATACCAAGATCATACAACAAGAGACTTTTCCCTCAAGGGGAGACT 2110
Qy 701 IleValGlnTyTrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Db 2111 ATAGTACAGTATACAGCACTGACCTGACCCCTCAGGCAATACAGAGCATGTGATATCATAT 2170
Qy 721 ValIleLysGlySerProCysGlnIleProPheThrProValAsnGlyAspPheIleCys 740
Db 2171 GTCTAATAGGCTCTCTCTGTGAATTCATTCACCTGTAAATGGGATTTTATATGC 2230
Qy 741 ThrProAspAsnThrGlyValAsnGlyThrLeuThrCysLeuGlnGlyTyTrAspPheThr 760
Db 2231 ACTCCAGATATATCTGAGTCACTGATATTAATCTGTGAGGGCTATGATTCACA 2290
Qy 761 GlnGlySerThrAspLysTyTrTyCysAlaTyTrGluAspGlyValITrpLysProThrTy 780
Db 2291 GAAGGGCTACTAGCAAGATATATGTGCTTATGAAAGATGGCGTGGAAACCAACATAT 2350
Qy 781 ThrThrGlnTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Db 2351 ACCACTAATGGCCAGACTGTGCCAAAACGTTTCTATACACGGGTTCAAGTCTTT 2410
Qy 801 GluMetPheTyTrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
Db 2411 GAGATGTTCTACAAAGCAGCTCGTGTGATGACACAGATCGATGAAGAAATTTTCTGAA 2470
Qy 821 AlaPheGlnThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
Db 2471 GCATTATAGACAGCCCTGGAAAAATGTCCTCAATTTGTGTATGATGACAGAGACAT 2530
Qy 841 AspCysArgLeuGlnGluAsnLeuThrLysLysTyTrCysLeuGlnTyTrAsnTyTrAspTy 860
Db 2531 GACTGCACTGAGAGAGAACTGACCAAAAATATGTCCTAGATATATATATATATAT 2590
Qy 861 GluAsnGlyPheAlaIleGlyProGlyGlyTyTrpGlyAlaAlaAsnArgLeuAspTySer 880
Db 2591 GAAAAATGGCTTGTGCAATTTGACCAAGTGGCTGGGGTCAAGCTATATAGGCTGATACCT 2650
Qy 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
Db 2651 TACATATACCTTCCGACACTGTGCAAGAAACACCAACCACTCGGAATGCCAAGTCC 2710
Qy 901 SerArgIleLysArgSerAlaProLeuSerAspTyTrLysIleLysLeuIlePheAsnIle 920
Db 2711 TCAGGATTAAGAAAGATGCCCATTTATCTGACTATTAATTAAGTTATTTTAACTTC 2770
Qy 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGlnTyTrpGluAsnGln 940
Db 2771 ACAGCTAGTGTGCATTACCCGATGAAGAAATGATACCTTGATGAAGGAAATCACAGAA 2830
Qy 941 ArgLeuLeuGlnThrLeuGlnTyTrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
Db 2831 CGACTCCTTCAAGCAATTTGAAACATATCACAAAATTAACGTGAAGAGACTCTCAACAAGAC 2890
Qy 961 ProMetTySerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGln 980
Db 2891 CCCATGATTCCTTTCAGCTTGCATCGAATTAATCTTATGCCGACAGCAATTCATTAGAA 2950
Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
Db 2951 ACAAAAAGGCTTCCCTCTGCAAGACAGGCTCAATGTGAGAGAGGCGATATGTGTTC 3010
Qy 1001 AsnCysProLeuGlyTyTrTyTrAsnLeuGlnHisPheThrCysGlyLysCysArgIle 1020
Db 3011 AATTGCCCTTTGGAGACTATTAATCTGAAACATTTCACTGTGAAAGTGTGCGGATC 3070
Qy 1021 GlySerTyTrGlnAspGlnGlnGlnLeuGlnCysLysLeuCysProSerGlyMetTy 1040
Db 3071 GATTCATATCAATGAAGAAAGGCACTTGAAGTGAAGCTTTCCTCTCGGATGTAC 3130
```

1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysValAlaGlnCysIleGlnGlyThr 1060
 Db AGGAATATATATCATTCATCAAGAAACATCTGATGTAACTCAGTGTAAACAAGGACCC 3190
 1061 TyrSerTyrSerGlnIleuGlnThrCysGlnSerCysProIleuGlnIleThrTyrGlnProIle 1080
 Db TACTCATACAGTGAAGCTGAGACTGTGAATCGGTCCAGCTGGCACTTATCAGCCAAA 3250
 1081 PheGlySerArgSerCysLeuSerCysProGlnIleAsnThrValIleValArgIleValAla 1100
 Db TTTGGTCCCGAGAGCTGCTCTGTGTCCAGAAACCTCACTGAGAAAAGAGAGCC 3310
 1101 ValAsnIleSerAlaCysGlnValProCysProGlnIleGlyIlePheSerArgSerGlyIleu 1120
 Db GTGAACATTTCTGCAATGAGATTCCTGTGTCCAGAAAGAAATCTGTGGCTTGGGGTGA 3370
 1121 MetProCysHisProCysProArgAspTyrTrpIleProAsnAlaGlyIleValAlaPheCys 1140
 Db ATGCTCTGTACCCCAATGCTCTGTGACTATTCACATATGCAAGGAAAGGCTTCTGC 3430
 1141 LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlnCys 1160
 Db CTGGCTGTCTCCCTTTATGGAATCAACCCATTCCTGTGTCCAGATCACAAGATGT 3490
 1161 SerSerPheSerSerThrPheSerAlaAlaGlnIleSerValValProAlaIleSerIleu 1180
 Db TCMAAGTTTATAGTTCACTTCTCAGCGGAGAGAAAGTGTGGTCCCTGCTCTCT 3550
 1181 GlyHisIleIleValArgHisGlnIleSerSerGlnValPheHisGlnCysPhePheAsn 1200
 Db GGACATATTTAAAGAGGCAATGAAATCAGATCAGGATTTTCCATGAAATCTCTTTAAC 3551
 1201 ProCysHisAsnSerGlyThrCysGlnGlnIleuGlyValArgIleTyrValCysLeuCysPro 1220
 Db CCTGCAACAATGAGTGAACCTCCAGCAACTTGGGGGTGTATTTGTTTCTCTGTCCA 3610
 1221 LeuGlyTyrThrGlyLeuIleCysGlnThrAspIleAspGlnCysSerProIleuProCys 1240
 Db CTGTGATATACAGGCTTAAAGTGTGAACAGACATCATGATGACAGCCCATGCTTGC 3671
 1241 LeuAsnAsnGlyValCysIleAspIleuValGlyGlnPheIleCysGlnCysProSerGly 1260
 Db CTCMAACAATGAGATTTGTAAAGACTAGTGGGAAATTCATTTGTAGTCCCATTCAGT 3731
 1261 TyrThrGlyValIleArgCysGlnIleAsnIleAsnGlnCysSerSerSerProCysLeuAsn 1280
 Db TACACAGGCTCAGGGGTGTGAAGAAATTAAGATGATGAGTCACTCTGTTTAAT 3791
 1281 IysGlyIleCysValAlaAspGlyValAlaGlyTyrArgCysThrCysValIysGlyPheVal 1300
 Db AAAGAAATCTGTGTGATGGTGTGGCTGCTATCGTGTGACATGTGTGAAGGATTTGTA 3851
 1301 GlyLeuHisCysGlnThrGlnValAlaAsnGlnCysGlnIleSerAsnProCysLeuAsnAla 1320
 Db GGCTTCATATGTGTAAACAAGAGTCAATGAATGCCAGCAAAACCCTTAATAAATGCA 3911
 1321 ValCysGlnAspGlnValGlyIlePheLeuCysIleCysProProGlyPheLeuGlyThr 1340
 Db GTCTGTGAAGACAGAGTGGGGATTTCTGTGCAATGCCCACTGATTTTGGGTACC 3971
 1341 ArgCysGlyValAsnValAspGlnCysLeuSerGlnProCysIleAsnGlnValAlaThrCys 1360
 Db CGATGTGAAGAAAGAACTCGATGATGTCTCAGTCAAGCCATGCAAAATGAGGATCACTGT 4031
 1361 IysAsnGlyValAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
 Db AAAGAGGGGTCCAATAGCTCAGATGCTGTGTGCACTGCTTCAACAGATCACTGT 4091
 4091 AAAGAGGGGTCCAATAGCTCAGATGCTGTGTGCACTGCTTCAACAGATCACTGT 4150
 1381 GluLeuAsnIleAsnGlnCysGlnIleSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
 Db GAATTCACATCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4151
 1401 GluLeuAsnSerTyrSerCysIleCysGlnProGlyPheSerGlnAlaArgCysGlnThr 1420

4211 GAATTAATTCATACAGTGTAAATGTCAGCCAGAGATTTTCAGGCAAAAGGTGGAACA 4270
 1421 GluGlnSerThrGlyPheAsnLeuAspPheGlnValSerGlyIleTyrGlyTyrValMet 1440
 Db GACAGTCTACAGGCTTTAACTGATTTGAAGTTTCTGCACTTATGATATGTCATG 4271
 1441 LeuAspGlyMetLeuProSerIleuHisAlaLeuThrCysThrPheThrMetIleSerSer 1460
 Db CTAGATGGCATGCTCCATCTCTCCATGCTTAACTGTAACTGTAACTCTGGATGAATCTCT 4331
 1461 AspAspPheLeuTyrGlyThrProIleSerTyrAlaValAlaAsnGlnIleSerAspAsnThr 1480
 Db GACACATGAACTATGGAACACCAATCTCTATCCAGATTTGAATACGGGAGGAGCAATGAC 4391
 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValIleuTyrValAsnGlyIleGlnIleThr 1500
 Db TTGCTCTGACTATTTAAACGCTGGGTTCTTTATGTAATGCAAGGAAATTAACA 4451
 1501 AsnCysProSerValAlaAspGlyArgTrpHisAlaAlaIleThrTrpThrSerAla 1520
 Db AACTGTCCCTCGGTAATGATGGCAGATGGCATCATATTGCAATCACTTGGACAAGTACC 4511
 1521 AsnGlyIleTrpIleValIleAspGlyIleLeuSerAspGlyIleGlyAlaGlyIleuSer 1540
 Db AATGGCATCTGGAAGCTATATCGATGGGAAATTAATCTGACGGTGTGCTGGCTCTCT 4571
 1541 ValGlyLeuProIleProGlyIleGlyAlaLeuValIleuGlyGlnGlnAlaAspIleIle 1560
 Db GTTGTGTGCTCCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4631
 1561 GlyIleGlyIlePheSerProAlaGlnIleSerPheValGlySerIleSerGlnLeuAsnLeuTrp 1580
 Db GGAGAGGATATCAGCCAGCTGAGTCTTTGTGGCTCCATTAACCACTGATCAACTCTGG 4691
 1581 AspTyrValLeuSerProGlnGlnValIleSerIleuAlaThrSerCysProGlnIleu 1600
 Db GACTATGTCTGTCTCCACAGCAGTGAAGTCACTGTGCTCTGCTGCCAGAGAACTC 4751
 1601 SerIleGlyAsnValIleuAlaTrpProAspPheLeuSerGlyIleValGlyIleValIle 1620
 Db AGTAAAGAAACGTGTTAAGATGCTGATTTCTGTCAAGAAATGTGGGAAATGTAAG 4811
 1621 IleAspSerIleSerIlePheCysSerAspCysProIleuGlnIleGlyIleSerValProHis 1640
 Db ATCGATTTCAAGACATATTTGTGTGATTTGCCACAGCTTAAAGAGGCTCAGTCTCAT 4871
 1641 LeuArgThrAlaSerGlnAspIleuValProGlySerIleValAlaLeuPheCysAspPro 1660
 Db CTAGAACTGCACTGAAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTCTGTGATCCA 4931
 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnIleGlnTrpThrGln 1680
 Db GGCTTCAGCTGTGGGAAACCTGTGCACTGTGTGATCAAGAGACAGTGAACAA 5051
 1681 ProLeuProHisCysGlnIleArgIleSerCysGlyValProProProIleuGlnIlePhe 1700
 Db CCACTTCTCAGTGAACCACTTAACTGAGGTGCACTCTGTGAGAAATGGCTTC 5051
 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValIleThrTyrGlnCysAsnAsnGlyTyr 1720
 Db CATTCACCCATGACTCTATGCTGTGAGAGCAAGTAACTTAACAGTGAACAAAGGCTAC 5111
 1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValIleSer 1740
 Db TATCTATTTGGGTGACTCAAGATGTTCTGTACAGTAAATGAGAGCTGGAACGGGTTTCA 5171
 1741 ProSerCysLeuAspValAlaAspGlnCysAlaValGlySerAspCysSerGlnIleAlaSer 1760
 Db CCATCTCGCTTATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5231
 1761 CysLeuAsnValAlaAspGlySerTyrIleCysSerCysValAlaProProTyrThrGlyAspGly 1780

481 GInG1yAenSerG1nTrpAspG1yProG1uProArgCyseValG1uArgHisCyseSerThr 500
 1451 CAAAGAAACAGCCAGTGGGATGGCCAAAGACCCGGTGTGTGGAGGCCACTGTTCCACC 1510
 501 PheG1metProLyseAspVal11e11eSerProHisAsnCyseG1yLyseG1nProAlaLyse 520
 1511 TTTCAAGATGCCCAAGATATATCATATATCCCCCAACACTGTGGCAAGCCAGCCCAAA 1570
 521 PheG1yThr11eCyseTyrValSerCyseArgG1nG1yPhe11eLeuSerG1yVal1yG1u 540
 1571 TTTGAGACATCTGCTATATAGTTCGCCCAAGGGGTTCATTTTATCTGGAGTCAAAGAA 1630
 541 MetLeuArgCyseThrThrSerG1yLyseTyrAsnValG1yValG1nAlaValCyseLyse 560
 1631 ATGCTGAGATGTACCACTTCTGGAAAATGGAAATGTGGAGTTCAAGCACTGTGTGAAA 1690
 561 AspValG1uAlaProG1n11eAsnCyseProLyseAsp11eG1uAlaLyseThrLeuG1n 580
 1691 GACGTGAGGGCTCTCAAAATCAACTGCTTAAGGACATAGAGCTTAAGACTCTGGAAACAG 1750
 581 G1nAspSerAlaAsnVal11eThrTrpG1n11eProThrAlaLyseAsnSerG1yG1uLyse 600
 1751 CAAGATCTGCAAGATGTAACCTGGCAGATTCCAACAGCTTAAGCAACTGTGGTGAAG 1810
 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPhePro11eG1yAspVal 620
 1811 GTGTCAAGTCACAGTTTCATCAAGCTTTCACCCCACTTACCTTTTCCAGTGGAGATGTT 1870
 621 Ala11eValTyrThrAlaThrAspLeuSerG1yAsnG1nAlaSerCyse11ePheHis11e 640
 1871 GGTATCGTATACACGGCACTGACTATCCGGCAACAGGCCCACTGGCAATTTTCCATATC 1930
 641 LysVal11eLeuAspAlaG1uProProVal11eAspTyrCyseArgSerProProProValG1n 660
 1931 AAGGTATTTGAGTACAGACCACTGTCTATAGACTGGTGGAGATCTCACTCCGTCAG 1990
 661 ValSerG1uLyseValHisAlaAserTyrAspG1uProG1nPheSerAspAsnSerG1y 680
 1991 GTCTCGAGAGAGTACATCCCGCACTGGAGATGAGCTTCAGATTCACAGCACTCAGGG 2050
 681 AlaG1uLeuVal11eThrArgSerHisThrG1nG1yAspLeuPheProG1nG1yG1uThr 700
 2051 GCTGAATGTGTATTAACCAAGATCATACCAAGAGACCTTTTCCCTCAAGGGAGACT 2110
 701 11eValG1nTyrThrAlaThrAspProSerG1yAsnAsnArgThrCyseAsp11eHis11e 720
 2111 ATAGTACGATATACACCACTGACCTCCCTCAGGCAATACAGACATGTATATCATATT 2170
 721 Val11eLyseG1ySerProCyseG1u11eProPheThrProValAsnG1yAspPhe11eCyse 740
 2171 GTCATTAAGAGTTCCTCCCTGTAATTCATTCACACTGTMAATGGGGATTTTATATGC 2230
 741 ThrProAspAsnThrG1yValAsnCyseThrLeuThrCyseLeuG1uG1yTyrAspPheThr 760
 2231 ACTCCAGATTAATACGTGAGTCAACTGTATTAATCTTGGAGGGCTATATGATTTTCACA 2290
 761 G1uG1ySerThrAspLyseTyrTyrCyseAlaTyrG1uAspG1yVal11eTyrLyseProThrTyr 780
 2291 GAAGGGTCTACTGACAGATATATGTGCTTATGAAAGTGGCCGTGGAAACCAATAT 2350
 781 ThrTrpG1uTyrProAspCyseAlaLyseLyseArgPheAlaAsnHisG1yPheLyseSerPhe 800
 2351 ACCACTGAATGGCCAGACTGTGCCAAAACCGTTTGTCTAACACCGGGTTCAGTCTCTT 2410
 801 G1uMetPheTyrLyseAlaAlaArgCyseAspAspThrAspLeuMetLyseLysePheSerG1u 820
 2411 GAGATGTCTTACAAAGACAGCTGTGTGTATGACACAGATCTGATGAAAGATTTTCTGAA 2470
 821 AlaPheG1uThrThrLeuG1yLyseMetValProSerPheCyseSerAspAlaG1uAsp11e 840
 2471 GCATTGACACACCTGGGAAAATGTGTCCATTCATTTTGTATGTATGACAGAGACATT 2530

841 AspCyseArgLeuG1uG1uAsnLeuThrLyseTyrCyseLeuG1uTyrAsnTyrAspTyr 860
 2531 GACTGCAGACTGAGAGAGAACTGACCAAAAATATATGCTTACAAATATATATATGACTAT 2590
 861 G1uAsnG1yPheAla11eG1yProG1yG1yTyrG1yAla11eAsnArgLeuAspTyrSer 880
 2591 GAAAATGTCTTGCATTAATGACAGAGGTGGCTGGGGTGCACCTTAATAGGCTGGATTACT 2650
 881 TyrAspAspPheLeuAspThrValG1nG1uThrAlaThrSer11eG1yAsnAlaLyseSer 900
 2651 TACGATGACTCTTCTGGACACTGTGCAAGAAACGCCCAACAGCATCGGCATTCGCAACTCC 2710
 901 SerArg11eLyseArgSerAlaProLeuSerAspTyr11eLyse11ePheAsn11e 920
 2711 TCAAGATTAATAAGAGTGCCTCATATATGATTAATAATTAATTAATTAATTAATTAATC 2770
 921 ThrAlaSerValProLeuProAspG1uArgAspAspThrLeuG1uTyrG1uAsnG1n 940
 2771 ACAGCTAGTGTGCATTACCCGATGAAAGAAATGATCCCTTGATGGGAAATCAGCAA 2830
 941 ArgLeuLeuG1nThrLeuG1uThr11eThrAsnLyseLeuLyseArgThrLeuAsnLyseAsp 960
 2831 CGACTCTTCAGACATTGAGAACTATTCACAAATTAACATGAAGAGACTTCAACAAAGAC 2890
 961 ProMetTyrSerPheG1nLeuAlaSerG1u11eLeu11eAlaAspSerAsnSerLeuG1u 980
 2891 CCCATGATATCTTTCAGCTGTGATCAAGAAATCTTATATACCCCAACAAATTCATTAAGAA 2950
 981 ThrLyseLyseAlaSerProPheCyseArgProG1ySerValLeuArgG1yArgMetCyseVal 1000
 2951 ACAAAAGAGCTTCCCTCTCTGACACAGGCTCGATGAGAGGAGGAGGAGGAGGAGGAGGAGG 3010
 1001 AsnCyseProLeuG1yThrTyrTyrAsnLeuG1uHisPheThrCyseG1uSerCyseArg11e 1020
 3011 AATTGCCCTTGGGAACATTAATTAATCTGGAACATTTCACTGTAAACCTGCCGATC 3070
 1021 G1ySerTyrG1nAspG1uG1yG1nLeuG1uCyseLyseCyseProSerG1yMetTyr 1040
 3071 GGAATCTATCAAGATTAAGAAAGGAGCACTTGATGCAAGCTTGGCTCCCTCGGATATAC 3130
 1041 ThrG1uTyr11eHisSerArgAsn11eSerAspCyseLyseAlaG1nCyseLyseG1nG1yThr 1060
 3131 ACAGAAATATATCAATTCAGAAACATCTCTGATTTGAAGCTCACTGTAACCAAGAGCAC 3190
 1061 TyrSerTyrSerG1yLeuG1uThrCyseG1uSerCyseProLeuG1yThrTyrG1nProLyse 1080
 3191 TACTCATAGAGTGAAGCTTGAAGCTTGAATCGTTCACCTGGGACATTAACGCCAAA 3250
 1081 PheG1ySerArgSerCyseLeuSerCyseProG1uAsnThrSerThrValLyseArgG1yAla 1100
 3251 TTTGGTTCCTCGAGGCTGCTCTCTGTGTCCGAAAACACTCACTGTGAAAAGAGAGGCC 3310
 1101 ValAsn11eSerAlaCyseG1yValProCyseProG1uG1yLysePheSerArgSerG1yLeu 1120
 3311 GTAAACATTTCTCAGATGAGAGTTCCTGTGTCAGAAAGAAAATTTCTGCTTGGGTTA 3370
 1121 MetProCyseHisProCyseProArgAspTyrTyrG1nProAlaAlaG1yLyseAlaPheCyse 1140
 3371 ATCCCGTCAACCAAGTCTCTGTGACTATTAACAACCTTAATCAGGAGGAGGAGGAGGAGGAGG 3430
 1141 LeuAlaCyseProPheTyrG1yThrThrProPheAlaG1ySerArgSer11eThrG1uCyse 1160
 3431 CTGGCCGTGCTCTTATATGAGACTACCCCAATGCGGTGTCAGATTCACATCAAGAAATGT 3490
 1161 SerSerPheSerSerThrPheSerAlaAlaG1uG1uSerValProProAlaSerLeu 1180
 3491 TCAAGTTTATGTTCAACTTCTCAGCGGCAAGAGAAAGTGTGTGCTCCCTGCTCTCTT 3550
 1181 G1yHis11eLyseLyseArgHisG1u11eSerSerG1nValPheHisG1uCysePhePheAsn 1200
 3551 GACATATATTAAGAGGAGCATGAATCAGACGTCAGGTTTTCATGATATGCTTCTTAAC 3610
 1201 ProCyseHisAsnSerG1yThrCyseG1nG1uLeuG1yArgG1yTyrValCyseLeuCysePro 1220

Db ||||| CCTCCCAAAATAGTGAACCTCCAGCAACTGGGGGTGTATGTTCTCTCTCCCA 3670
 Qy 1221 LeuGlyTyrThrGlyValLeuValCysGlyLeuThrAspIleAspGluCysSerProLeuProCys 1240
 Db 3671 CTGGATTAATACAGGCTTAAAGTGTGAACAGACATCATGATGACGCCCACTGCTTGC 3730
 Qy 1241 LeuAsnAsnGlyValAlaCysIleAspLeuValGlyIlePheIleCysGluCysProSerGly 1260
 Db 3731 CTCAACAATGAGACTTGTAAAGACCTAGTTGGGGAATTCATTTGTAGTCCCATCAAGT 3790
 Qy 1261 TyrThrGlyValIleCysGlyValIleAsnGluCysSerSerProCysLeuAsn 1280
 Db 3791 TACACAGGTCAAGGGGTGTGAAGAAATATTAATGAGGTAGCTCCAGTCTCTGTTTAAAT 3850
 Qy 1281 LysGlyIleCysValAlaAspGlyValAlaGlyTyrThrCysThrCysValLysGlyPheVal 1300
 Db 3851 AAGGAAATCTGTGTATAGGTGTGGCTGTCTATCTGTGCAATGTGTGAAGAAAGATTGTA 3910
 Qy 1301 GlyLeuHisCysGluThrGluValAlaAsnGluCysGlnSerAsnProCysLeuAsnAsnAla 1320
 Db 3911 GGCTGCAATGTGTGAACAGAGTCAATGATGCAAGTCAACCAATGCTTAATTAATGCA 3970
 Qy 1321 ValCysGluAspGlnValGlyIlePheLeuCysValCysProProGlyPheLeuGlyThr 1340
 Db 3971 GTCTGTGAAGACACAGGTTGGGGGATTTCTGTGCAAAATGCCACCTGGATTTTGGGTACC 4030
 Qy 1341 ArgCysGlyLysAsnValAspGluCysLeuSerGluProCysLysAsnGlyValAlaThrCys 1360
 Db 4031 CCAATGTGAAGAAACCTCGATGAGTGTCTGACGACGACATGCAAAATGAGCACTACTGT 4090
 Qy 1361 LysAspGlyValAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys 1380
 Db 4091 AAGAGCGTCCCAATGCTTCAATGCTGTGTGCAAGTGGCTTCAAGAGTCAACTGT 4150
 Qy 1381 GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp 1400
 Db 4151 GAATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 4210
 Qy 1401 GluLeuAsnSerCysSerCysGlnProGlyPheSerGlyGlnAspCysGluThr 1420
 Db 4211 GAATTAACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 4270
 Qy 1421 GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet 1440
 Db 4271 GAAAGCTCAACAGGCTTAAACCTGGAATTTTGAAGTTTCTGGCATCATGATATGTGATG 4330
 Qy 1441 LeuAspGlyMetLeuProSerIleuHisAlaLeuThrCysThrPheThrMetLysSerSer 1460
 Db 4331 CTAGATGGCATGTCTCCATCTCTCCATGCTCTAAACCTGTAACCTTCTGGAATGAATCCTCT 4390
 Qy 1461 AsnAspMetAsnThrGlyThrProLysSerTyrAlaValAspAsnGlySerAspAsnThr 1480
 Db 4391 GACGATATTAATCATATGAAACCAATCTCTATGCAAGTGTAAACGACGACGCAATATCC 4450
 Qy 1481 LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyValGlyLysIleThr 1500
 Db 4451 TTGCTCTCACTGATTAATTAACGCTGGCTTCTTATGTGAAATGCGAGGAAAGATTAACA 4510
 Qy 1501 AsnCysProSerValAsnAspGlyValArgTyrPheHisIleAlaIleThrTyrThrSerAla 1520
 Db 4511 AACTGTCCCTCGGTGATATGATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 4570
 Qy 1521 AsnGlyIleTyrPheValIleTyrIleAspGlyLysLeuSerAspGlyValIleGlyLysSer 1540
 Db 4571 AATGGATCTGGAAGTCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4630
 Qy 1541 ValGlyLeuProIleProGlyGlyGlyAlaLeuValLeuGlyGlnGluAsnLysLys 1560
 Db 4631 GTTGGTTTGGCCCATACCTGT 4690
 Qy 1561 GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGluLeuAsnLeuTyr 1580

Db 4691 GAGAGGGATTCAAGCCAGCTAGTCTTTGTGGGCTTCATTAACCACTCAACTCTGG 4750
 Qy 1581 AspTyrValLeuSerProGlnIleValLysSerLeuAlaThrSerCysProGluGlyLeu 1600
 Db 4751 GACTATGCTCTGTCTCCACAGCAGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4810
 Qy 1601 SerIleGlyAsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyLysValLys 1620
 Db 4811 AGTAAGAAACCGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4870
 Qy 1621 IleAspSerLysSerIlePheCysSerAspCysProGluGlyGlySerValProHis 1640
 Db 4871 ATGATATCTAAGACATATTTTGTGATTTGCTTGAATGCTTGAAGAGGCTGATCTCAT 4930
 Qy 1641 LeuArgThrAlaSerGluAspLeuValProGlySerLysValAsnLeuPheCysAspPro 1660
 Db 4931 CTGGAACCTGCACTGAAATTTTAACCAAGCTTCCAAAGTCAATCTGTTCTGTGATCCA 4990
 Qy 1661 GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGln 1680
 Db 4991 GGCTTCCAGCTGGTGGGAAACCTGTGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 5050
 Qy 1681 ProLeuProHisCysGluValGlyIleSerCysGlyValAlaProProProLeuGluAsnGlyPhe 1700
 Db 5051 CCACTTCTCACTGATGAACCAATTAAGCTGTGGGTGTCACCTCTTGTGAAGATGGCTTC 5110
 Qy 1701 HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
 Db 5111 CATTCAAGCCATGATCTTCTATGCTGCGACAGACATTAACATCAAGTCAACAAATGGCTAC 5170
 Qy 1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrPasnGlyValSer 1740
 Db 5171 TATCTATTTGGTGACTCAAGATGTTCTGTACAGATTAATGAGACCTGAAACGGGCTTCA 5230
 Qy 1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
 Db 5231 CCATCTGCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5290
 Qy 1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
 Db 5291 TGCCTGAACCTAATGATGATCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5350
 Qy 1781 LysAsnCysAlaGluProIleLysCysValAlaProGlyAsnProGluAsnGlyHisSer 1800
 Db 5351 AAAAAGCTGACAGAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5410
 Qy 1801 SerGlyGluIleTyrThrValGlyAlaValAlaThrPheSerCysGlnGluGlyTyrGln 1820
 Db 5411 TCAAGTGAATTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5470
 Qy 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrPasnHisLeuIlePro 1840
 Db 5471 TTGATGGAGATTAACCAAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5530
 Qy 1841 TyrCys 1842
 Db 5531 TATTGT 5536

RESULT 3

AX686198

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AX686198 11152 bp DNA linear PAT 29-MAR-2003
 Sequence 7 from Patent WO20059315.
 AX686198.1 GI:29372029
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 Shinkete, R.A., Patturajan, M., Vermont, C.A., Gasman, S.J.,
 Malyankar, V., Shenoy, S., Spytek, K.A., Gangolli, B., Miller, C.,
 Boldog, F., Li, L., Taupier, R.J., Kerkuda, R., Smithson, G.,

Zerhoben, B.D., Liu, X., Colman, S.D., Tchernev, V., Si, J., Edinger, S.,
Stone, D., Schiore, P., Miller, I. and Rothenberg, M.
Human nucleic acids and polypeptides and methods of use thereof
Patent: WO 02059315-A 7 01-AUG-2002;

FEATRES
source

Location/Qualifiers
1. 11152
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	11152
Score:	9862.00	Matches:	1781
Percent Similarity:	97.72%	Conservative:	16
Best Local Similarity:	96.85%	Mismatches:	40
Query Match:	96.51%	Indels:	2
DB:	6	Gaps:	1

US-09-977-053-6 (1-1842) x AX686198 (1-11152)

```
Oy      4 ArgLeuAlaPheCysCysTrrGlyLeuAlaLeuValSerGlyYTrpAlaThrPheGlnGln 23
Db      83 AGAATTTCGCGCGCTTGTCTGGGGTCTGGGCTCGTTTCGGGCTGGGCGACCTTTCAGCG 142
Oy     24 MetSerProSerArgAspPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db     143 ATGTCCCGTCCGCGAATTCAGCTTCGCCCTTCCCGAAGCCGCGCCCGGGCCCCC 202
Oy     44 GlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
Db     203 GCGAGTATCCCGCGCGCGCGCTCTGGCGAAGAGCGCGGGAGAGAGTGAAGCGG 262
Oy     64 LeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
Db     263 CTGGGCGCGCGCTTC-----CGCGCGCGCTGCGCGGAGCTCAGCGAGCGCTGGAG 316
Oy     84 LeuValPheLeuValAspAspSerSerValGlyGluValAspPheArgSerGluLeu 103
Db     317 CTGTCTTCTCTGTGGATGATTCGTCCAGCGTGGCGAGAGTCACTTCGCGAGCGAGCTC 376
Oy    104 MetPheValArgLeuLeuLeuSerAspPheProValProThrAlaThrArgValAla 123
Db     377 ATGTTCGCGCGCGAAGCTGTCCAGCTTCCCGTGGTCCCAAGCGCGCGCGTGGCC 436
Oy    124 IleValThrPheSerSerIleAspThrValProArgValAspThrIleSerThrArg 143
Db     437 ATCGTACCTTCTGTCCAGAACTACGTGTGCGCGCGCTGATTAACATCTCCACCCGC 496
Oy    144 ArgAlaArgGlnIleCysAlaLeuLeuLeuGlnIleProAlaIleSerIleArg 163
Db     497 CGCGCGCGCGCGAAGTGGCGCGCTGCTCTCCAGAGATCCCTGCCATCTCCAGCA 556
Oy    164 GlyGlyGlyThrTrpThrIleGlyAlaPheGlnGlnAlaGlnIleLeuLeuHisAla 183
Db     557 GGTGCGCGCACTTACCAAGGCGCGCTTCCAGAAAGCGCGCAATCTTCTTCAGCT 616
Oy    184 ArgGluAsnSerThrIleValValPheLeuIleThrAspGlyTrpSerAsnGlyGlyAsp 203
Db     617 AGAGAAACCTCAACAAAGTGTATTTCTCATCACTGATGATATTCATAGGGAGAGC 676
Oy    204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db     677 CTTAGACCAATTTGCGAGCTGCTGCGAGATTCAGAGATGAGATCTTTCATTTGGCATA 736
Oy    224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProGlyGluGluHisCys 243
Db     737 TGGCAAGGAGCAATTCGAGAGCTGATGACATGGCTTCCACCCCAAGAGAGAGCACTGT 796
Oy    244 TyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeuHisGluAsp 263
Db     797 TACCTGTACACAGTTTGAAGAAATTTGAGGCTTTAGCTGCGCGGCAATTGATGAGAT 856
```

```
Oy    264 LeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrpLeuCysAspGlu 283
Db     857 CTACCTTCGGAGAGTTTATTCAGAGTATGATGTCATCGCTCATATCTTGTGATGAG 916
Oy    284 GlyIleAspCysCysAspArgMetGlySerCysIleCysGlyThrIleThrGlyHisPhe 303
Db     917 GCGAAGGACGTGCTGAGCCGAATGGAGAGCTCCAAATGGAGACACACAGGCGCATTTT 976
Oy    304 GlyCysIleCysGluIleGlyTrpTrpGlyIleGluIleThrGluCysThrAlaCys 323
Db     977 GAGTGCATCTGTGAAGAGGGATTTACGGGAAGGCTGCAATGATATGCACGCTTGC 1036
Oy    324 ProSerGlyThrTrpLeuProGluGlySerProGlyGlyIleSerSerCysIleProCys 343
Db    1037 CCATCGGGAGACATACAAACCTCAAGAGCTTCAAGAGAGATCAGAGTTGCATTCATGT 1096
Oy    344 ProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCysValCysArg 363
Db    1097 CCGATGAAATCACACCTCTCCACTGGAGACATCCCTGAAAGACTGTGTCTCAGA 1156
Oy    364 GlyGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLysPro 383
Db    1157 AGAGATACAGGCGCATCTGGCGAGCTGTGAATCTTTCATCTGCTGCCCTGCGAAGCT 1216
Oy    384 ProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHisPheAlaAlaCysGly 403
Db    1217 CCGAATAAGTTACTTATTCAAACACTTCACAAACCACTTCATCAATCAGCCTGTGG 1276
Oy    404 ValArgCysHisPheProGlyPheAspLeuValGlySerSerIleIleLeuCysIleProAsn 423
Db    1277 GTCCAGATGACACCTGGATTTATCTTGGAGAGAGATCATCTTATGTCTTACCAAT 1336
Oy    424 GlyLeuTrpSerGlySerGluSerTrpCysArgValArgThrCysProHisLeuArgGln 443
Db    1337 GGTTCGTGTCCGGTTCAAGAGCTTACTCAGAGTAAAGACATGTCTCATCTCCCGCAG 1396
Oy    444 ProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTrpLysThrCysLeu 463
Db    1397 CCGAAACATGGCCACATCAGCTGTCTTACAGGAAATGTATATGACACATCTTTG 1456
Oy    464 ValAlaCysAspGluGlyTrpArgLeuGluGlySerAspLysLeuThrCysGlnIleAsn 483
Db    1457 GTTGCCTGTGAAGAGGTACAGACTGAAGAGCAGATGATGATCTTACTGCAAGAAAC 1516
Oy    484 SerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGlnMet 503
Db    1517 AGCCAGTGGAGTGGCGCAGAACCCCGGTGTGTGAGCGCCACTGTTCACCTTCAGATG 1576
Oy    504 ProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPheGlyThr 523
Db    1577 CCGAAGAGTATCATATATCCCCCAACACTGTGGCAACAGCAACCAATTTGGAGAG 1636
Oy    524 IleCysTrpValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMetLeuArg 543
Db    1637 ATCTGTATGATGATGTCCGCAAGGTTCAATTTATCTGAGTCCAAAGAAATGCTGAGG 1696
Oy    544 CysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLysAspValGlu 563
Db    1697 TGTACCACTTCTGGAAGAAATGGAATGTCCGAGTTCAAGCACTGTGTGTAAAGCGTGAAG 1756
Oy    564 AlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGlnIleAspSer 583
Db    1757 GCTCCCAATATCACTGTCTTAAAGCATAAGAGGCTTAAACACTCTGAAACAGCAAGATTTCT 1816
Oy    584 AlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyLysLysValSerVal 603
Db    1817 GCCAATGTTACCTGTGAGATTCACAGAGCTTAAAGAACTCTGTGTAAAGAGGTGTACGTC 1876
Oy    604 HisValHisProAlaPheThrProProTrpLeuPheProIleGlyAspValAlaIleVal 623
Db    1877 CGGTTTCATCCAGCTTTCACCCCACTTATCTTTCCAAATTTGAGATGTGTCTACGTA 1936
```

```

Qy 624 TyrThrAlaThrAspLeuSerGlyValSerGlnAlaSerCysIlePheHisIleValIle 643
Db 1937 TACACGGCAACATCACTATCCGGCAACACAGGCCAGCTGTCATTTTCATATCAAGGTTAT 1996
Qy 644 AspAlaGluProProValIleLeuSerProCysArgSerProProValGlnValIleSerGln 663
Db 1997 GATGCAAGAACCACTGTCATAGACGTGGTGAAGATCTCCACCTCCGTCAGGTCCTGAG 2056
Qy 664 LysValHisAlaAlaIleSerTrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeu 683
Db 2057 AAGTACATATGCCCAAGCTGGATGAGCCTCAGTTCACAGCAACTCAGGGGGCTGAATTG 2116
Qy 684 ValIleThrArgSerHisIleThrGlnGlyAspLeuPheProGlnGlyIleValGln 703
Db 2117 GTCATTACCAAGAGATATACACAGAGAGACCTTTCCCTCAAGGGAGACATATAGTACAG 2176
Qy 704 TyrThrAlaThrAspProSerGlyValAsnAsnArgThrCysAspIleHisIleValIleLys 723
Db 2177 TATACAGCACTGACCCCTCAGGGTATACAGAGATATGATATCCATATTCATATGCAATAA 2236
Qy 724 GlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCysThrProAsp 743
Db 2237 GGTTCCTCCCTGTAATTCATTCACACTCTGTAATGGGATTTTATATGCACTCCAGAT 2296
Qy 744 AsnThrGlyValAsnCysThrLeuThrCysLeuGlnGlyTyrAspPheThrGlnGlySer 763
Db 2297 AATACAGAGTCAACTGATCACTTAACTTGGAGGGCTACAGATTCACAGAAAGGCTCT 2356
Qy 764 ThrAspLysTyrTyrCysAlaTyrGlnAspGlyValIleTrpLysProThrTyrThrGln 783
Db 2357 ACTGACAGATATATATGCTTATGAGAGATGGCGTCCGAAACCAACATATACCACTGAA 2416
Qy 784 TrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGluMetPhe 803
Db 2417 TGGCCAGACTGTCGCCAAAACCTTTTGCAACACCGGTTCAAGTCTTTGAGATGTTCT 2476
Qy 804 TyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlnAlaPheGln 823
Db 2477 TACAAAGCAGCTCGTGTGATGACTCAGATCTGATGAGAAAGTTTCTGAAAGCATTTGAG 2536
Qy 824 ThrThrLeuGlnLysMetValProSerPheCysSerAspAlaGluAspIleAspCysArg 843
Db 2537 ACGACCTTGAGAAATATGGTCCATCATTTTGTAGATGCAAGAGACATATGACTGACAG 2596
Qy 844 LeuGlnGluAsnLeuThrLysLysTyrCysLeuGlnTyrAsnTyrAspTyrGlnAsnGly 863
Db 2597 CTGAGAGAAACCTGACCAAAAATTTGCTAGATATATATGACTATGAAATGGC 2656
Qy 864 PheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAsp 883
Db 2657 TTTGCAATTTGCTCAGATGGCTGGGTCGAGCTAATATGGCTGATTAATCTTACGATGAC 2716
Qy 884 PheLeuAspThrValGlnGlnThrAlaIleThrSerIleGlyValAsnAlaLysSerSerArgIle 903
Db 2717 TTCCTGAGACACTGTGCAAGAAACAGCACAGATGGGCAATGCCAAGTCTCAAGGAT 2776
Qy 904 LysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSer 923
Db 2777 AAAAGAGGCCCTCATATCTGACTATTAATTAAGTTAATTTTAACTACACAGCTAGT 2836
Qy 924 ValProLeuProAspGlyLysArgAsnAspThrLeuGlnTyrGlnAsnGlnArgLeuLeu 943
Db 2837 GAGCCATTTCCGATGAAAGAAATGATACCTTGATATGGGAAATATGAGCAACACATCTCT 2896
Qy 944 GlnThrLeuGlnThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyr 963
Db 2897 CAGACATTTGAAACATATCAAAATTAACGTGAAAGAGCTCTCAACAAAGCCCAAGTAT 2956
Qy 964 SerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlnThrLysLys 983
Db 2957 TCCTTTACGCTTGACATCAAAATTAATTAAGCCGACAGCAATTCATTTAGAAACAAAAG 3016
Qy 984 AlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysPro 1003

```

```

Db 3017 GCTTCCCCCTTCGACACACAGGCTCAGTGTGAGAGGCCATATGTGTCATATGGCCCT 3076
Qy 1004 LeuGlyThrTyrTrpAsnLeuGlnHisPheThrCysGluSerCysArgIleGlySerTyr 1023
Db 3077 TTGGAACTTATATATATCTGAAACATTTCACTCTGTAAACCTGCGGATGGATCTAT 3136
Qy 1024 GlnAspGluGlnGlyLeuGlnCysLysLeuCysProSerGlyMetTyrThrGluTyr 1043
Db 3137 CAAGATGAAGAGGCAACTTGAAGTGAAGCTTTGCCCTCTGGAGATGACACGGAAATAT 3196
Qy 1044 IleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyr 1063
Db 3197 ATTCATTCAGAAACATCTCTGATTTGTAAGCTCAGTGAACAAAGGACCTACATCAATC 3256
Qy 1064 SerGlyLeuGlnThrCysGluSerCysProLeuGlnTyrTyrGlnProLysPheGlySer 1083
Db 3257 AGTGGACTTGAGACTTGATATGCTGTCACATGGGCACTTATCAGCCAAATTTGGTTCC 3316
Qy 1084 ArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIle 1103
Db 3317 CAGAGCTGCCCTCTGTCTCAGAAACACCTCACTGGAAGAAAGAGCCGTGAACAT 3376
Qy 1104 SerAlaCysGlyValProCysProGlnGlyLysPheSerArgSerGlyLeuMetProCys 1123
Db 3377 TCTGCATGTGAGTATCTCTGTCCAGAGAAATTCGCGCTTCGGGTATATGCCCTGT 3436
Qy 1124 HisProCysProAspArgTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCys 1143
Db 3437 CACCATGCTCTGAGCTATATACCACTTATACAGAGAAAGCTTCTGCTGGCTGT 3496
Qy 1144 ProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlnCysSerSerPhe 1163
Db 3497 CCGTTTATGAACTACCCCATTCGCTGCTGCCAGATCCATCAAGATGTTCAAGTTT 3556
Qy 1164 SerSerThrPheSerAlaAlaGluGlnSerValValProProAlaSerLeuGlyHisIle 1183
Db 3557 AGTTCAACTTCTCAGGGGAGAGAAAGTGTGTGCCCTGCTCTCTTGGACATAT 3616
Qy 1184 LysLysArgHisGlnIleSerSerGlnValPheHisGlnCysPhePheAsnProCysHis 1203
Db 3617 AAAAGAGCAATGATATCAGCAGTCAAGCAAGTCAATATATGCTTATACCTTGGCAC 3676
Qy 1204 AsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyr 1223
Db 3677 AATATGGAACCTGCCAGCAACTTGGCGGTATATTTGTCTGTCCACTTGGANAT 3736
Qy 1224 ThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsn 1243
Db 3737 ACAGGTTTAAAGTGTGAACAGACATCGATGAGTGCAGGCCCACTGCTTGCCTCAACAT 3796
Qy 1244 GlyValCysLysAspLeuValGlyGluPheIleCysGlnCysProSerGlyTyrThrGly 1263
Db 3797 GAGGTTTGTAAGACCTAGTGGGAAATTCATTTGTAGAGGCCCATCATGCTTACACAGGT 3856
Qy 1264 GlnArgCysGlnLeuAsnIleAsnGlnCysSerSerSerProCysLeuAsnLysGlyIle 1283
Db 3857 AAGCAGCTGATATGAACATCAATGAATGTCAGTCAATCATATGATGAATCAAGCCACC 3916
Qy 1284 CysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGlyLeuHis 1303
Db 3917 TGTGTGATTAATTAATTCATACAGTGTGAATGTTCAGCCAGGATTTTCAGGCAAAAG 3976
Qy 1304 CysGlnThrGlnValAsnGlnCysGlnSerAsnProCysLeuAsnAlaValCysGln 1323
Db 3977 TGTGAACAGTATGATATCACTCAGTGTATTAATTAACCTTATATGCACTGTGAA 4036
Qy 1324 AspGlnValGlyIlePheLeuCysLysCysProProGlyPheLeuGlnTyrArgCysGly 1343
Db 4037 GACCAAGTGGGAGATCTTGTGCAAAATGCCCACTGGAATTTTGGGTACCCGATGCGA 4096
Qy 1344 LysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAspGly 1363

```

Db 4097 AAGAGCTGATGATGCTCAGTCAGCAGCATGCAAAAATGAGCTACCTGTAAGACGGT 4156
 Qy 1364 AAlaanserPheargCysleuValAlaAlaGlyPheThrGlySerHisCysGluLeuLeu 1383
 Db 4157 GCCAAATGCTCAGAGCTGCTGTGTGAGCTGCTTCAAGAGATCACTGGAATTGAAC 4216
 Qy 1384 ILaasngluCysGlnSerAspProCysArgAsnGlnAlaThrCysValAspGluLeuLeu 1403
 Db 4217 ATCAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4276
 Qy 1404 SerTyrSerCysGlyCysGlnProGlyPheSerGlyGlnArgGluThrGluGlnSer 1423
 Db 4277 TCATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4336
 Qy 1424 ThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuAspGly 1443
 Db 4337 ACAGGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4396
 Qy 1444 MetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLeuSerSerAspAspMet 1463
 Db 4397 ATGCTCCATCTCTCCATGCTCTTACCTGATCTTCTGATGATGATGATGATGATGAT 4456
 Qy 1464 AsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeuLeu 1483
 Db 4457 AACTATGAGAACCAATCTCTATGATGATGATGATGATGATGATGATGATGATGATGAT 4516
 Qy 1484 ThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyTyrGluGluGluGluGluGlu 1503
 Db 4517 ACTGATTTAAACGGGTGGGTCTTTATGATGATGATGATGATGATGATGATGATGATGAT 4576
 Qy 1504 SerValAspAspGlyArgTyrHisIleAlaIleThrTyrThrSerAlaAsnGlyLe 1523
 Db 4577 TGGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4636
 Qy 1524 TrpValValTyrIleAspGlyTyrLeuSerAspGlyGlyValGlyLeuSerValGlyLeu 1543
 Db 4637 TGGAGGCTAT 4696
 Qy 1544 ProIleProGlyGlyValAlaLeuValLeuGlyGluGluAspTyrValGlyGlyGly 1563
 Db 4697 GCCAATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4756
 Qy 1564 PheSerProAlaGlySerPheValGlySerIleSerGluLeuAsnLeuTyrAspTyrVal 1583
 Db 4757 TTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4816
 Qy 1584 LeuSerProGlnGlnValIleSerLeuAlaThrSerCysProGluGluLeuSerValGly 1603
 Db 4817 CTGTCTCCACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4876
 Qy 1604 AsnValLeuAlaTyrProAspPheLeuSerGlyIleValGlyValIleAspSer 1623
 Db 4877 AACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4936
 Qy 1624 LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr 1643
 Db 4937 AAGAGATATTTTGTCTGATTTCCACGCTTGGAGGGGCTCATGCTCATCGAAGACT 4996
 Qy 1644 AlaSerGluAspLeuLysProGlySerIleValAsnLeuPheCysAspProGlyPheGln 1663
 Db 4997 GCATCTGAAGATTAAACCAAGGTTCCAAAGTCAATCTGTTGTTGTTGTTGTTGTTG 5056
 Qy 1664 LeuValGlyAsnProValGlnTyrCysLeuAsnGluGlyGluThrThrGlnProLeuPro 1683
 Db 5057 CTGGTGGGAAACCTGTGAGTACTGTGTAATCAAGAGAGTGAACAACAACCTCC 5116
 Qy 1684 HisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHisSerAla 1703
 Db 5117 CACTGGAACGATTTGGCTGTGGGGTCCACTCTCTTGGAGATGCTTCATTCAGCC 5176
 Qy 1704 AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrTyrLeu 1723
 Db 5177 GATGACTTATCTGCTGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5236

Qy 1724 GlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerProSerCys 1743
 Db 5237 GGTACTCTAAGAGATGTTCTGTGACAGATTAATGGAGCTGGAACGGGGTTCACCATCTG 5296
 Qy 1744 LeuAspValAspGluCysValAlaGlySerAspCysSerGluHisAlaSerCysLeuAsn 1763
 Db 5297 TTAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5356
 Qy 1764 ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnGly 1783
 Db 5357 GTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5416
 Qy 1784 AlaGluProIleLysCysValAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu 1803
 Db 5417 GCAGAACTTATTAATGTAAGCTCCAGAAATCCGGAATATGCGCACTCTCCAGGTAG 5476
 Qy 1804 IleTyrThrValGlyValAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly 1823
 Db 5477 ATTATTAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5536
 Qy 1824 ValThrValIleThrCysLeuGluGluSerGlyGluTyrAsnHisAlaIleProTyrCys 1842
 Db 5537 GTACCAAAATCAATGTTTGGAGCTGAGAGATGGAATCATCTATATATATATATATAT 5593

RESULT 4

AX686196 11158 bp DNA linear PAT 29-MAR-2003
 LOCUS Sequence 5 from Patent WO02059315.
 DEFINITION AX686196
 ACCESSION AX686196.1 GI:29372028
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Shinkens, R.A., Patterson, M., Vernet, C.A., Casman, S.J.,
 Malpankar, U., Shenoy, S., Szytek, K.A., Gangoli, B., Miller, C.,
 Boldog, F., Li, L., Taupier, R.J., Kekuda, R., Smithson, G.,
 Zerhusen, B.D., Liu, X., Coleman, S.D., Tcherny, V., St. J., Edinger, S.,
 Stone, D., Sciore, P., Miller, I. and Rothenberg, M.
 Human nuclear acids and polypeptides and methods of use thereof
 Patent: WO 02059315-A 5 01-AUG-2002;
 Curesgen Corporation (US)
 Location/Qualifiers

FEATURES

source 1..11158
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 11158
 Score: 9812.00 Matches: 1774
 Percent Similarity: 97.23% Conservative: 16
 Best Local Similarity: 96.36% Mismatches: 47
 Query Match: 96.02% Indels: 4
 Gaps: 2

US-09-977-053-6 (1-1842) x AX686196 (1-11158)

Qy 4 ArgLeuAlaPheCysCysTyrProGlyLeuAlaLeuValSerGlyTyrAlaThrPheGlnGln 23
 Db 83 AGAATTGGCGGCTTGTCTGGGCTTGGGCTGCTGCTTCTGGGCTGAGACCTTTCAGCAG 142
 Qy 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyValAPro 43
 Db 143 ATGTCCTCCGTCGCGCAATTCAGCTTCGCTCTTCCCGAAGACCGGCGGCGGCGG 202
 Qy 44 GlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArgValGluArg 63
 Db 203 GGAAGATATCCCGCGCGCGCGCTCTCTGCGAGAGACGGCGGCGGAGACAGTGAACGG 262

QY 64 LeuGIyGIAlaPheArgValArgLeuArgGluLeuSerGIuArgLeuGIu 83
 DB 263 CCGGGCCAGGCGTTC-----CGCGTCGCGCTCGGGAGGCTCAGCGAGCGCGTGGAG 316
 QY 84 LeuValPheLeuValAspAspSerSerValGlyValAspPheArgSerGIuLeu 103
 DB 317 CTGTCTCTCTGGTGGATGATTCGTCCAGCGTGGCCAAAGTCACTCCCGCAGGAGCTC 376
 QY 104 MetPheValArgPheLeuSerAspPheProValValProThrAlaThrArgValAla 123
 DB 377 ATGTGTCTCGGCAAGCTGCTGCTCGACTCCCGTGTGCGCCAGGCCAGCGCGTGGCC 436
 QY 124 IleValThrPheSerSerValAsnValValProArgValAspValIleSerThrArg 143
 DB 437 ATGTGACCTTCTGCTCAAGACTACGTGTGCGCGCGCTGATTCATCTCCACCGC 496
 QY 144 ArgAlaArgGlnIleValValAlaLeuLeuGlnGluIleProAlaIleSerThrArg 163
 DB 497 CGCGCGCGCAGCAAGTGCCTGCTCTCCAGAGATCCCTGCGCATCTCTACCGA 556
 QY 164 GlyGIyGIyThrThrThrArgValAlaPheGlnGlnAlaGlnIleLeuLeuHla 183
 DB 557 GGTGGCGGACCTACACCAAGGCGCTCTCCAGAGCGCGCAAAATCTCTTCATGCT 616
 QY 184 ArgGluAsnSerThrValValPheLeuIleThrAspGIyTrSerAsnGIyAsp 203
 DB 617 AGAGAAACCTCAACAAAGTGTATTTCTCATCACTATGATGATTCATTCAGGGGAGAC 676
 QY 204 ProArgProIleAlaAlaSerLeuArgAspSerGIyValGluIlePheThrPheGIyIle 223
 DB 677 CCTAGACCAATTCGAGGCTCATGCGAGATTCAGAGATGAGATCTTCATCTTGGCA 736
 QY 224 TrpGIyGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProGlyGluGluHla 243
 DB 737 TGGCAAGGAACTTCAGAGCTCATGATGATGCTTCACCCCAAGAGAGACACTGT 796
 QY 244 TyrLeuLeuHlaSerPheGlnGluPheGluAlaLeuAlaArg-----ArgAlaLeuHla 261
 DB 797 TACTGTGACACGATTTGAAAGATTTGAGGCTTTAGCGCCCTGTCATATGATATTT 856
 QY 262 GluAspLeuProSerGIySerPheIleGlnAspAspMetValHlaSerThrLeuArg 281
 DB 857 GTAGATCTACCTCTGAGAGATTTATTCAGATGATATGTCACCTGCTCATCTTGT 916
 QY 282 AspGluGlyAspAspAspMetGlySerCysLeuGlyThrHlaSerThrGIy 301
 DB 917 GATGAAGGCAAGACTGCTGACCAATGAGAGCTGCAAAATGGGAAACACAGAGC 976
 QY 302 HisPheGluCysIleCysGluGlyGlyTyrTyrGIyGlyGlyLeuGlnTyrGluCysThr 321
 DB 977 CATTTGAGTGCATCTGTGAAAGGGGATTAACGGGAAAGGTCTGCAGTATGACTGCA 1036
 QY 322 AlaCysProSerGIyThrTyrLeuProGluGlySerProGIyGlyIleSerSerCysIle 341
 DB 1037 GTTGGCCCATCGGGGACATCAAACTGAAGGCTCACACAGAGAAATCAGAGTTCATT 1096
 QY 342 ProCysProAspGluAsnHlaSerSerProGIySerThrSerProGluAspCysVal 361
 DB 1097 CATGTCCTATGAAATCACTCTTCACCTGGAAGCAATCCCTGAAAGACTGTGTC 1156
 QY 362 CysArgGluGlyTyrArgAlaSerGIyGlnThrCysGluLeuValHlaCysProAlaLeu 381
 DB 1157 TGCAGAGAGGATTCAGGGGATTCGCGCAGACTGTGAAGTTGTCCACTGCGCCCTG 1216
 QY 382 LeuProProGluLeuGlyTyrPheIleGlnAsnThrCysAsnAsnHlaSerAspAlaAla 401
 DB 1217 AACCCCTCCGAAATGATCTTATCCAAACCTGCAACCACTTCATTAAGCAGCC 1276
 QY 402 CysGIyValArgCysHlaSerProGlyPheAspLeuValGlySerSerIleIleLeuCysLeu 421
 DB 1277 TGTGGGTCGCAATGTCACCTGAGATTTGATCTTGTGGAGACAGATCATCTTATGCTTA 1336

QY 422 ProAsnGlyLeuTrpSerGIySerGluSerTyrCysArgValArgThrCysProHlaSer 441
 DB 1337 CCGAATGCTTGTGTGTCGCGTTCAGAGAGCTACAGAGATGAAGAACATGCTCATCTC 1396
 QY 442 ArgGlnProLeuHlaGlyHlaIleSerCysSerThrArgGluMetLeuTyrLeuThr 461
 DB 1397 CGCAGCCGAAACATGCGCACHATGCTCTTCACAGGAAATTTATTAAGCAACA 1456
 QY 462 CysLeuValAlaCysAspGluGlyTyrArgLeuGlnGlySerAspLeuThrCysGln 481
 DB 1457 TGTGTGTGCTGTGATGAAGGATGACAGCTAGAGAGCAATGATTAACCTTACTGCA 1516
 QY 482 GlyAsnSerGlnTrpAspGIyProGluProArgCysValGluArgHlaCysSerThrPhe 501
 DB 1517 GGAACAGCCAGTGTGATGGCGCAAGACCCGCGTGTGAGCGCATGTTCCACTTT 1576
 QY 502 GlnMetProLeuAspValIleIleSerProHlaAsnCysGlyLeuGlnProAlaSerPhe 521
 DB 1577 CAGATGCCCAAGATGTCATATATCCCAACACTGTGGCAAGCACCCCAAAATTT 1636
 QY 522 GlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGIyValIysGlnMet 541
 DB 1637 GGAAGATCTGCTATGTAAGTTGCGCCAGCGTTCACTTATCTGAGTCAAAAGAAATG 1696
 QY 542 LeuArgCysThrThrSerGIyGlyGlyTrpAsnValGlyValGlnAlaValCysIysAsp 561
 DB 1697 CTGAGATGTACACTCTCGAAATGGAATGCGAGTTCAGGCGCGCTGTGTAAAGC 1756
 QY 562 ValGluAlaProGlnIleAsnCysProLeuAspIleGluAlaLeuThrLeuGluGln 581
 DB 1757 GTGAGGCTCTCAATATCACTGCTTACAGACATGAGGCTAGACACTTGGAAACGCA 1816
 QY 582 AspSerAlaAsnValThrTrpGlnIleProThrAlaLeuAspAsnSerGIyGlyValVal 601
 DB 1817 GATTCGCAATGTACTGCTGCAATTCACAGCTTAAGCAACTCTGTGTAAGAGTG 1876
 QY 602 SerValHlaSerValHlaSerProAlaPheThrProProTyrLeuPheProIleGlyAspValAla 621
 DB 1877 TCAGTCCACGTTTACCTCAGACTTTCACCCCACTTATCCCAATTTGAGATGTGCT 1936
 QY 622 IleValTyrThrAlaThrAspLeuSerGIyAsnGlnAlaSerCysIlePheHlaIleLeu 641
 DB 1937 ATGCTATACAGGCACTGACCTATCCGCAACAGGCAAGCTGATTTCCATTCAG 1996
 QY 642 ValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGlnVal 661
 DB 1997 GTTATGATGCAAGAACCACTGTATGACATGATGCAATCTCCCGCTCAGAGTC 2056
 QY 662 SerGluValHlaAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGIyAla 681
 DB 2057 TCGAGAGAGGTACATGCGCAGAGCTGGGATGAGCTCAGTTCACAGAACTCAGGTGCT 2116
 QY 682 GluLeuValIleIleThrArgSerHlaSerThrGlnGlyAspLeuPheProGlnGlyIle 701
 DB 2117 GAAATTTGCTATTCAGAGATCAATACAGAGAGACTTTTCTCCAGAGGAGACTATA 2176
 QY 702 ValGlnTyrThrAlaThrAspProSerGIyAsnAsnArgThrCysAspIleHlaIleVal 721
 DB 2177 GTACAGTATACAGCACAGCACTGACCCCTCAGGCAATACAGCACTGATATATCAATTC 2236
 QY 722 IleLeuGlySerProCysGluGluIleProPheThrProValIleGlnGlyAspPheIleCysThr 741
 DB 2237 ATAAAGGTTCTCCCTGGAATTCATTCACCTGTAAATGGGAAATTTATTAATGACT 2296
 QY 742 ProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
 DB 2297 CCAATATATCTGAGATCACTGATCAATTAATCTGAGAGGCTATGATTCACAGAA 2356
 QY 762 GlySerThrAspValTyrTyrCysAlaTyrGluAspGIyValTrpLeuProThrTyrThr 781
 DB 2357 GGGTCTACTGACAGATATATGCTTATGAAGATGCGCTGGAACCAACATATATCC 2416
 QY 782 ThrGluTrpProAspCysAlaLeuValArgPheAlaAsnHlaGlyPheLeuSerPheGlu 801

Db 2417 ACTGATGCGCAGACGTGCTCAGTAAGCGCTTTTGCAAAACGAGGCTTCAAGTCTCTTGAG 2476
 Qy MetPheTyTylValAlaArgCysAspAspThrAspLeuMetClyAspPheSerGluAla 821
 Db 2477 AGTTTCTCAAGAGCGCTGTTGTGTATGACACAGATCTGATGAAAGATTCTTCTGAGCA 2536
 Qy PheGluThrThrLeuGluTyrMetValProSerPheCysSerAspAlaGluAspIleAsp 841
 Db 2537 TTGAGACGACCCCTGGGAAAATGCTCCCATCTTTGTGATGATCAGAGGACATTGAC 2596
 Qy CysArgLeuGluGluAsnLeuThrTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
 Db 2597 TGCAGCTGAGAGAGACCTGACCAAAAATATATGCTGAAATATATATATATATGAA 2656
 Qy AsnGlyPheAlaIleGlyProGlyValTyrGlyValAlaAsnArgLeuAspTyrSerTyr 881
 Db 2657 AATGCGCTTGGCAATGTCCTCAGGTGCTGGGTGCGATTAATGAGCTGATTAATCTTAC 2716
 Qy AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaIleYssSer 901
 Db 2717 GATGACTTCTGACACTGTGCAAGAAACAGCCACAGATCGGCAATCGCAATGCTCTCA 2776
 Qy ArgIleTyrArgSerAlaProLeuSerAspTyrTyrIleTyrLeuIlePheAsnIleThr 921
 Db 2777 CGGATTAAGAAAGAGTCCCATATATCTGACTAATAATTAAGTTAATTTTAAATCATACA 2836
 Qy AlaSerValProLeuProAspGluAlaGlnAspThrLeuGluTyrProIleAsnGlnAlaArg 941
 Db 2837 GCTAGTGTGCGCATTCACCGATGAAAGAAATGATATACCTTGAAATGGGAAATACGCAACA 2896
 Qy LeuLeuGlnThrLeuGluTyrIleThrAsnTyrLeuTyrArgThrLeuAsnTyrAspPro 961
 Db 2897 CTCCTTCAGACATTTGAAACTATACAAATTAATTAAGAAAGACTTTCACAAAGAGCCCC 2956
 Qy MetTyrSerPheGluLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluTyr 981
 Db 2957 ATGTAATCTCTTCAGCTTCATCAAGAAATTAATTAATGAGCGAATTCATTAAGAAACA 3016
 Qy LysValAlaSerProPheCysArgProGlySerValLeuArgIleTyrMetCysValAsn 1001
 Db 3017 AAAAAGGCTTCCCTCTTGCAGACAGGCTCAGTCTGAGAGGCGGTATGTTGTCAT 3076
 Qy CysProLeuGluTyrTyrTyrAsnLeuGluIlePheThrCysGluSerCysArgIleGly 1021
 Db 3077 TCCCTTTGGGAACCTATTAATCTGAAACATTTCACTGTGAAGCTGCGGATCGGA 3136
 Qy SerTyrGlnAspGluGluIleGlnLeuGluCysValLeuCysProSerGlyMetTyrThr 1041
 Db 3137 TCCATCAAGATGAAGAGGCACTTGAGTGCMAAGCTTGGCCCTCGGGAATGTACAG 3196
 Qy GluTyrTyrIleAsnSerArgAsnIleSerAspCysValAlaGlnCysValGlnGluTyr 1061
 Db 3197 GAATATATCATTCAGAAACATCTGTATGTAAAGCTCAGTGAACAGGACACTTAC 3256
 Qy SerTyrSerGlyLeuGluTyrCysGluSerCysProLeuGluTyrTyrGlnProIlePhe 1081
 Db 3257 TCATACAGTGAACCTTGACCTGTGATGTGTCCACTGGGCACTTAATACGCAAAATTT 3316
 Qy GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValIleValAlaVal 1101
 Db 3317 GGTGCCCCGAGCGCTCTCGTGTCCAGAAACACCTCAACTGTGAAGAGAGGACCGTG 3376
 Qy AsnIleSerAlaCysGlyValAlaProCysProGluGluTyrPheSerArgSerGlyLeuMet 1121
 Db 3377 AACATTTCTGCAATGTGAGATCTCTGTGCCAAGAGAAATCTCGGCTCTGGGTTAAAG 3436
 Qy ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyIleValAlaPheCysLeu 1141
 Db 3437 CCTCTGACCCCATGTCTCTGTGACTATTAACCACTTAATGCAAGGAGGAGGCTTCTGCG 3496
 Qy AlaCysProPheTyrGluTyrThrProPheAlaGlySerArgSerIleThrGluCysSer 1161

Db 3497 GCGTGTCCCTTTATGAGAACTACCCCATTCGCTGTTCCAGATCATACAGAAATTTCA 3556
 Qy SerPheSerSerThrPheSerAlaIleGluGluSerValIleProProAlaSerLeuGly 1181
 Db 3557 AGTTTATGTTCAACTTTCACAGGCGAGAGAAAGTGTGTCCCTGCTCTTGGGA 3616
 Qy HisIleTyrAsnArgHisGluIleSerSerGlnValPheHisGluCysPhePheAspPro 1201
 Db 3617 CATATTAAGAAAGGCAATGAATCAGACAGTCAAGGCAAGTCAATGAATGCTTCTTAAACCT 3676
 Qy CysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeu 1221
 Db 3677 TGCACAAATAGTGAACCTGCGCAGCAACTGGGCGGTATATGTGTCTCTGTCCACTT 3736
 Qy GlyTyrThrGluLeuLeuSerGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
 Db 3737 GGAATATACAGGTTTAAGGTGAAGACATGATGAGTGCAGGCCCATGCTGCTTC 3796
 Qy AsnAsnGlyValCysIleAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
 Db 3797 AACATGAGATTGTAAAGACCTAGTTGGGAAATTCATTTGTGAGTCCCATAGGTTAC 3856
 Qy ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnTyr 1281
 Db 3857 ACAGTGAAGCACTGTGAATGGAACATCAATGAATGTCAGTCTAATCCATGTGAATCAG 3916
 Qy GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValIleGlyPheValGly 1301
 Db 3917 GCCACTGTGTGAATTAATTAATCAATACAGTTGTAATGTGACGCAAGATTTTCAGGC 3976
 Qy LeuHisCysGluThrGluValAsnGluCysGlnSerAspProCysLeuAsnAlaVal 1321
 Db 3977 AAAAGGTGTGAACAGATGATGATCAATCAAGTATTAATACTTAATATATGACGTC 4036
 Qy CysGluAspGlnValGlyGlyPheLeuCysValCysProProGlyPheLeuGlyThrArg 1341
 Db 4037 TGTGAAGACACAGGTTGGGGATTCCTGTGCATAATGCCACCTGGAATTTTGGTACCCGA 4096
 Qy CysGlyTyrAsnValAspGluCysLeuSerGlnProCysValAsnGlyAlaThrCysVal 1361
 Db 4097 TGTGAAGAACCTCAATGATGTCTCAGTCAAGCCATGCAAAATGAGACTTACTGTAAA 4156
 Qy AspGlyAlaAsnSerPheArgCysLeuCysValAlaGlyPheThrIleCysGlu 1381
 Db 4157 GACGGTGCCAATAGCTTCAAGTGCCTGTGTGACGCTGCTCAACAGATCACTGTGAA 4216
 Qy LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
 Db 4217 TTGAACATCAATGAATGTCACTTAATCAATGAAGAAATCAGGCTGTGTGATGAA 4276
 Qy LeuAsnSerTyrSerCysValCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
 Db 4277 TTAAATTCATACAGTTGTAATGTCAAGCAAGATTTTCAAGGCAAAAGGTGTGAACAGAA 4336
 Qy GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
 Db 4337 CAGTTCACAGGCTTAACTGGAATTTGAAGTTTGGGATCTAATGATATGTCAAGCTA 4396
 Qy AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetIleYssSerAsp 1461
 Db 4397 GATGGATGCTCCCATCTCCATAGCTTAACCTGTGACTTCTGGAATGAATCTCTGAC 4456
 Qy AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlyYssAspAsnThrLeu 1481
 Db 4457 GACATGAACCTATGAAACCAATCTCTATGAGAGTTGATACGAGGACCAATACCTTG 4516
 Qy LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyTyrGlyValIleThrAsn 1501
 Db 4517 CTCCTGACGATTAATTAACGAGGTGCTTATGTGAATGCGAGGAAAGATTAACAAAC 4576
 Qy CysProSerValAsnAspGlyArgGlyPheHisIleAlaIleThrThrThrSerAlaAsn 1521
 Db 4577 TGTCTCGGTATATGATGCAAGATGCAATATTAATGCAATCACTTGACAAATGACTGCT 4636

OY	1522	GlyIleTrrPlyValIYrIleApgIyLysLeuSerApgIyValAgiLysLeuSerVal	1541	ACCESSION	AF206329
DB	4637	GGAGCCTGGAGGGCTATATAATGGGGGATTAATCTGACGGTGTACTGGCCTCTCATTT	4696	VERSION	AF206329.1
OY	1542	GlyLeuProIleProGlyGlyValAlaLeuValLeuGlyValGlnGlyAspIlyLeuGly	1561	KEYWORDS	GI:11177163
DB	4697	GGGAAAGCCATACCTGCTGGCGGCTGATTAAGTCTTGGGCGCAAGGCAAGCAAAAAAGAA	4756	SOURCE	Mus musculus (house mouse)
OY	1562	GluGlyPheSerProAlaGlySerPheValGlySerIleSerGlnLeuLeuTrrPasp	1581	ORGANISM	Mus musculus
DB	4757	GAAGGATTCAACCCGCTGAGTCTTTTGTGGCTCCATTAAGCCAGCTCAACCTCTGGGAC	4816	REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OY	1582	TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGlnGlnLeuSer	1601	AUTHORS	Gallego, D., Vinit, M.A., Callebaut, I., Coulombel, L., Cacheux, V., Romeo, P.H. and Vigon, I.
DB	4817	TATGTCCTGCTCCACGACGAGTGAAGTCACTGGCTACTCTGCTACCTGCGACAGAACTCAGT	4876	TITLE	Polydom: a secreted protein with pentraxin, complement control protein, epidermal growth factor and von Willebrand factor A domains
OY	1602	LysGlyAsnValLeuAlaTrrProAspPheLeuSerGlyIleValGlyValLeuIle	1621	JOURNAL	Biochem. J. 352 Pt 1, 49-59 (2000)
DB	4877	AAAGGAAACGTGTACATGGCTGATTTCTTGTCAAGAAATGTGGGAAAGTGAAGATC	4936	MEDLINE	20517255
OY	1622	AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHileu	1641	PUBMED	11062057
DB	4937	GATTCTAAGACATATTTTCTGATTTGCCCAAGCTTGGGAGGGTCAAGTCCCTCATCTG	4996	REFERENCES	2 (bases 1 to 11289)
OY	1642	ArgThrAlaSerGlnAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly	1661	AUTHORS	Gallego, D., Vinit, M.-A., Callebaut, I., Coulombel, L., Cacheux, V., Romeo, P.-H. and Vigon, I.
DB	4997	AGAACTGCATCTGAAGTTTAAACCAAGGTTCCAAAGTCAATCTGTTCTGAAACCAAGCC	5056	TITLE	Direct Submission
OY	1662	PheGlnLeuValGlyAsnProValAlaGlnTrrCysLeuAsnGlnGlyValTrrPheGlnPro	1681	JOURNAL	Submitted (18-NOV-1999) u474, Inserm, Maternite Port-Royal, 123 bid de Port-Royal, Paris 75014, France
DB	5057	TTCCACTGCTGGCGGAACCTGTGCACTGCTGAATCAAGACAGTGAACAACA	5116	FEATURES	Location/Qualifiers
OY	1682	LeuProHileCysGlyLysArgIleSerCysGlyValProProLeuGlnAsnGlyPheHile	1701	SOURCE	1..11289
DB	5117	CTCCCACTGTAACCACTTCGCTGTGGGGTGCACCTCTTTGGAAATGGCTTCAT	5176	CDS	1..11289
OY	1702	SerAlaAspAspPheTrrAlaGlySerThrValThrTrrGlnCysAsnAsnGlyTrrTrr	1721	codon_start=1	/product="polydom protein"
DB	5177	TCAGCCGATGACTTCTATGCTGGCAGACAGTAACCTACCAAGTGCACAAATGGCTCAT	5236	/protein_id="AA032160.1"	/db_xref="GI:11177164"
OY	1722	LeuLeuGlyAspSerArgPheCysThrAspAsnGlySerTrrPheAsnGlyValSerPro	1741	/translation="MSEKLPCCNALALVSGMTNPQVAYASLNSFLPPEASGALCRLAVPPASSEBAAGSRVERLGRFNRVERLRLSLSLVPLVDSSSVGGQNFALBLKRVRLSLDFPVVSTATVAIVTSSKNVAVRVYISTSRARHOCALLSRIPA	/db_xref="taxon:10090"
DB	5237	CTATTGGGTGACTCAAGGATGTTCTGTACAGATTAATGGAGCTGGAACGGGGTTTACCA	5296	/chromosome="4"	/map="4b-4c2"
OY	1742	SerCysLeuAspValAspGlnCysAlaValGlySerAspCysSerGlnHileAlaSerCys	1761	/cell_line="MS-5 stromal cells"	221..10924
DB	5297	TCTGCTTGAATGTCATGAGTGTGAGTGGATCAATGTGATGAGCANTGCTTCTTGC	5356	/codon_start=1	/product="polydom protein"
OY	1762	LeuAsnValaAspGlySerTrrIleCysSerCysValProProTrrPheGlyLys	1781	/protein_id="AA032160.1"	/db_xref="GI:11177164"
DB	5357	CTGAACGTATAGATGATCTTACATATATGTTCAATGTCTCCACCGTACACAGAGGATGGAAA	5416	/translation="MSEKLPCCNALALVSGMTNPQVAYASLNSFLPPEASGALCRLAVPPASSEBAAGSRVERLGRFNRVERLRLSLSLVPLVDSSSVGGQNFALBLKRVRLSLDFPVVSTATVAIVTSSKNVAVRVYISTSRARHOCALLSRIPA	/db_xref="taxon:10090"
OY	1782	AsnCysAlaGluProIleLysCysGlyValaProGlyAsnProGlyLysLeuSerSer	1801	/db_xref="taxon:10090"	/map="4b-4c2"
DB	5417	AACTGTGCAAACTATAAATGTAAAGCTCCAGGAAATCCGGAATAATGCGACTCTCTCA	5476	/chromosome="4"	/map="4b-4c2"
OY	1802	GlyGlnIleTrrThrValGlyAlaAlaValThrPheSerCysGlnGlnGlyTrrGlnLeu	1821	/cell_line="MS-5 stromal cells"	221..10924
DB	5477	GGTGAAGTTTATACAGTAGTGTCCGAAGTACATTTTCTGTCAGAGAGATACCACTTG	5536	/codon_start=1	/product="polydom protein"
OY	1822	MetGlyValaThrLysIleThrCysLeuGlnGlySerGlyLysTrrPheAsnHileuIleProTrr	1841	/protein_id="AA032160.1"	/db_xref="GI:11177164"
DB	5537	ATGGAGATTAACAAATACATGTTTGAAGTCTGGAATGGAATCATCTAATACATAT	5596	/translation="MSEKLPCCNALALVSGMTNPQVAYASLNSFLPPEASGALCRLAVPPASSEBAAGSRVERLGRFNRVERLRLSLSLVPLVDSSSVGGQNFALBLKRVRLSLDFPVVSTATVAIVTSSKNVAVRVYISTSRARHOCALLSRIPA	/db_xref="taxon:10090"
OY	1842	Cys 1842		/db_xref="taxon:10090"	/map="4b-4c2"
DB	5597	TGT 5599		/chromosome="4"	/map="4b-4c2"
RESULT 5	AF206329	11289 bp	mus musculus polydom protein mRNA, complete cds.		
LOCUS	AF206329	11289 bp	mus musculus polydom protein mRNA, complete cds.		
DEFINITION	mus musculus polydom protein mRNA, complete cds.				


```

Db      3635  TCCCTGCTGCTGCTCCCTTTATGGAACCAACCATCATCGCGCCACGTCCTCAACAGAC 3694
Qy      1160  CysSerSerPheSerSerThrpheSerAlaIagIuSerValAlaProProIaIaSer 1179
Db      3695  TCTCAAGATTGATGCTTACTTCTCAGCAGAGAAAGAAAGCAATAGTGGCTCGTGGCC 3754
Qy      1180  LeuGIyHieIleIyLeuYargHiegiuIeSerSerGIuValPheHiegiuCyPhePhe 1199
Db      3755  CCTGACATCTCCAGAACAGTACAGAAAGTCAAGAGTCAAGGTCTTTTCAAGAACTCTTCA 3814
Qy      1200  AsnProCyVHieAsnSerGIyThrCySgInGIuLeuGIyValGIyValCySleuCyS 1219
Db      3815  AACCCCTGCAACACGTCGAACTGCAACAGCTTGGCGGTGATGCTGCTCTCTG 3874
Qy      1220  ProIeuGIyTyrThrGIyLeuYsCySgIuThrAspIleAspGIuCySerProIeuPro 1239
Db      3875  CCACTGATGATACAGGCTTAAAGTGTGAACAGATATGATGAATGACAGCTCTGCT 3934
Qy      1240  CySleuAsnAsnGIyValCySlyAspLeuValGIyGIuPheIleCySgIuCySProSer 1259
Db      3935  TCCCTCAATGGTGAATTTGTAGAGCAAGTTGGGGGATTCACGTCGCAATGTTCAATTG 3994
Qy      1260  GIYTYrThrGIyGIuArgCySgIuGIuAsnIleAsnGIuCySereSerSerProCySleu 1279
Db      3995  GGCATTTCAGTCAATATGATGAAGAAATATTAATGAGTATCTTCACGCTTCTTA 4054
Qy      1280  AsnIySgIyIleCyValAspGIyValIagIyTyrArgCySereThrCyValIySgIyPhe 1299
Db      4055  AATAAAGAAACCTGCACCTACCGCTTGGAGCTGACCGCTGACCTGTGTGAAGAGATAC 4114
Qy      1300  ValGIyLeuHieCySgIuThrGIuValIaAsnGIuCySgInSerAsnProCySleuAsn 1319
Db      4115  ATGGGTGTCACCTGTGAACAGACGTCATGATGACATGCAAGCCCTGCTTAAACAAC 4174
Qy      1320  AlaValCySgIuAspGIuValGIyGIyPheLeuCySlyCySProProGIyPheLeuGIy 1339
Db      4175  GCAGTTTGTAAAGAACCAAGTGGGGGTCTCATGCAAAATGCCACCGAGATTTTGGGT 4234
Qy      1340  ThrArgCySgIyLyAsnValAspGIuCySleuSerGIuProCySlyAsnGIyAlaThr 1359
Db      4235  ACTCGGTGTGAAGAAATATGATGATGATGATCTCAGTCAAGCAGTCAAGAAATGAGCACT 4294
Qy      1360  CySlyAspGIyAlaAsnSerPheArgCySleuCySAlaIagIyPheThrGIySerHie 1379
Db      4295  TGTAAAGATGGTCCAAACAGCTTCAAGTGTCAATGTCCAGCAGGCTTCAACAGGACAC 4354
Qy      1380  CySgIuLeuAsnIleAsnGIuCySgInSerAsnProCySlyArgAsnGIuAlaThrCySVal 1399
Db      4355  TGTGAACCTGAACATCAACAGATGTCAGTCAACCCATGTAGAACCAAGCCACTGTGAG 4414
Qy      1400  AspGIuLeuAsnSerTyrSerCySlySgInProGIyPheSerGIyGIuArgCySgIu 1419
Db      4415  GATGAACCTAACTCATACAGCTGTATATGTCAAGCCAGATTTTCAAGCCACAGGTGTGAG 4474
Qy      1420  ThrGIuInSerThrGIyPheAsnLeuAspPheGIuValSerGIyIleTyrGIyVal 1439
Db      4475  ACAGAAACGCTTCGCTTTTAACTGGATTTTGAAGTTTCTGCGCATCTCAAGGATGCTC 4534
Qy      1440  MetLeuAspGIyMetLeuProSerLeuHieAlaIeUThrCySereThrPheTyrPheSer 1459
Db      4535  CTGCTAGAGAGAGTGTCCCAACCCCTCAATGCCATTAACCTGCCATTCGAGAGAAATCC 4594
Qy      1460  SerAspAspMetAsnTyrGIyThrProIleSerTyrAlaValAspAsnIySerAspAsn 1479
Db      4595  TTTGATGTATCATCATACAGGAGCCCATCTCTAGTGTGACTTGAAGATGAACAAGCAAC 4654
Qy      1480  ThrLeuLeuLeuThrAspTyrArgGIyTyrValIeTyrValAsnGIyArgGIyIyIle 1499
Db      4655  ACCTCCCTCTCTAGTATCAACAACGCGCTGCTTATATGTGAATGAAAGAAAGATC 4714
Qy      1500  ThrAsnCySProSerValAsnAspGIyArgTyrHieIleIleAlaIeThrThrThrSer 1519

```

```

Db      4715  ACCAATGCCCCCTCCGTAATGATGCGATTTGGCATCATATTTGCCAATCATATGCAACAT 4774
Qy      1520  AlaAsnGIyIleTyrLySValTyrIleAspGIyLySleuSerAspGIyAlaIagIyLeu 1539
Db      4775  ACTGTGAGACCTCGAAGGCTTATATTAATGGGGAATTTATCGACCGTGTACTGGCTTC 4834
Qy      1540  SerValGIyLeuProIleProGIyGIyGIyAlaLeuValIeugIyGIuInAspLyS 1559
Db      4835  TCCATTGGCAAAAGCCATACCTGTGCGGTGCATTTGTGTCTTGGGCAAGAGCAAGCAAA 4894
Qy      1560  LySGIyGIuGIyPheSerProIaIagIuSerPheValGIySerIleSerGIuLeuAsnLeu 1579
Db      4895  AAAGAGAGAGGATTAACCGGCTGAGTCTTTGTGGCTTCAATAGCCAGCTCAACCTC 4954
Qy      1580  TrpAspTyrValIeSerProGIuInValIySereLeuAlaThrSerCySProGIuGIu 1599
Db      4955  TGGGATATGTCTGTCTCCACAGCAGTGAAGTTGCTGGCCAGCTCCGCCAGAGAA 5014
Qy      1600  LeuSerLySGIyAsnValIeuaIaTyrProAspPheLeuSerGIyIleValGIyLySVal 1619
Db      5015  CTGAGTGGGGAAACGTGTAGCATGCGCGATTTCTGTGCGGAATCACGGGAAAGTG 5074
Qy      1620  LySIIeAspSerLySereIlePheCySereAspCySProArgLeuGIyGIySerValPro 1639
Db      5075  AAGGTGATTTCAAGCAGCATGTTCTGCTGATGTCCGCTTTAAGAGATCCGTGCT 5134
Qy      1640  HieLeuArgThrAlaSerGIuAspLeuLySProGIySerLySValAsnIyPheCySAsp 1659
Db      5135  CACCTGAGACCTGCATCAAGAAATGAAACCAAGCTCCAAAGTCAAGTCTGTGAT 5194
Qy      1660  ProGIyPheGIuLeuValGIyAsnProValGIyTyrCySleuAsnGIuGIyInTyrThr 1679
Db      5195  CCGGGCTTCAGATGTTGGGAATCTGTGCAATATGCTGAACCAAGGCGAGTGACA 5254
Qy      1680  GIuProLeuProHieCySgIuArgIleSerCySgIyValProProProLeuGIuAsnGIy 1699
Db      5255  CAACCACTCCCTCAGTGAACCATTCGCTGTGGCTGCTCCCGCTTGAAGATGGC 5314
Qy      1700  PheHieSerAlaAspAspPheTyrAlaIagIySerThrValThrTyrGIuCySAsnAsnGIy 1719
Db      5315  TTCTACTCAGCCGAGACCTTCATGCGGACAGCAGCTGATCATGATGACCAAGTGGC 5374
Qy      1720  TyrTyrLeuLeuGIyAspSerArgPheCysGlyThrAspAsnGIySerTrpAsnGIyVal 1739
Db      5375  TACTACCTGCTGGTGAATTCGGAATGTTCTGCAAGAAACAGGAGCTGGAACGGGATTT 5434
Qy      1740  SerProSerCySleuAspValAspGIuCySAlaIaGIySerAspCySereGIuHieAla 1759
Db      5435  TACCATCTGCTGATGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 5494
Qy      1760  SerCySleuAsnValAspGIySerTyrIleCySereCySValProProTyrThrGIyAsp 1779
Db      5495  TCTGCTGTAACCAACCAACGAGATCTAGATATGTTCTGTAAACCAACATACAGGGAGAT 5554
Qy      1780  GIyLySAsnCyAlaIagIuProIleLySGIyLySAlaProGIyAsnProGIuAsnGIyHie 1799
Db      5555  GGGAAAACTGTGCAAGACCTGTAAATGAATGAAGCTCCAGAAATCCAGAAATGGCAC 5614
Qy      1800  SerSerGIyGIuIleTyrThrValGIyAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 1819
Db      5615  TCTTCTGTGAAGATTTACACGTTGAGTACTGCACTCATTTCTCTGTACGAAGGGCAC 5674
Qy      1820  GIuLeuMetGIyValThrLySIIeThrCySleuGIuSerGIyGIuThrPheHieIle 1839
Db      5675  GAGCTGTGGAGTGAACATCATCATGTGTTGGAGACTGCGAGTGGAGTCCCTCAGG 5734
Qy      1840  ProTyrCyS 1842
Db      5735  CGGTCTGT 5743

```

RESULT 6
AR435510
LOCUS AR435510 11230 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 3 from patent US 6656707.
 ACCESSION AR435510
 VERSION AR435510.1 GI:40198413
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 11230)
 AUTHORS Welch, A.A. and Billiot, G.S.
 TITLE C3b/C4b complement receptor-like molecules and uses thereof
 JOURNAL Patent: US 6656707-A 3 02-DEC-2003;
 FEATURES
 source 1..11230
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	11230
Score:	8652.50	Matches:	1532
Percent Similarity:	90.56%	Conservative:	137
Best Local Similarity:	83.13%	Mismatches:	171
Query Match:	84.67%	Indels:	3
DB:	6	Gaps:	3

US-09-977-053-6 (1-1842) x AR435510 (1-11230)

1 MetTrpProArgLeuAlaPheCySyrTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
 176 ATGTGTCGGCGCTGCTTTTGTGTGCTGGGCTGGACACTGGTGGGCTGGACCAAC 235
 21 PheGlnGlnMetSerProSerArgAspPheSerPheArgLeuPheProGlnThrAlaPro 40
 236 TTCACGCGCCCTGGCCCTTCGCTCACTCACTGCTCCGCTGTCCTCCGAGGCTCTCCG 295
 41 GlyAlaProGlySerLeuProAlaProAlaProGly---AspGlnAlaAlaGlySer 59
 296 GGGGCTCTGGGCAAGACTGGCGGTACTCCGCGTCCAGTGAAGAGAGGAGGAGGAGAC 355
 60 ArgValGlnArgLeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGlyLeuSer 79
 356 AAGTGAAGCGCTCGGCGCGCGGCTTCGGAGCGCGGTGGCGGACCTGGCGGACTCAGC 415
 80 GlnArgLeuGlnLeuValPheLeuValAspAspSerSerValGlyGlnValAspPhe 99
 416 GGCAGGCTGAGGCTCTGCTTCCTGTGTGAGACAGTCTCCAGCGTGGCCAAACCACTTC 475
 100 ArgSerGlnLeuMetPheValArgValLeuLeuSerAspPheProValValProArgValAspTyr 119
 476 CTCACGAGCTCAAGTTCGTGCGGACCTGCTGCCACTTCCCGGTGTCTCAACGGCC 535
 120 ThrArgValAlaLeuValThrPheSerSerLyAspAntyValValProArgValAspTyr 139
 536 ACCGGTGTGGCCACTGCTCACTTCTCATCAAGAAACAAGTGGTGGCGCGGTGGATTAC 595
 140 IleSerThrArgArgAlaArgGlnIleGlyGlyAlaLeuLeuLeuGlnGlnIleProAla 159
 596 ATCTCCACACGCGCGCGGACCAACCAAGTGCCTGCTCAACCGCCGAGATCCCGGCC 655
 160 IleSerThrArgArgGlyGlyThrTyrThrTyrGlyGlyAlaPheGlnGlnAlaAlaGlnIle 179
 656 ATCACTTACCGCGGTGTGGCACTTATCCAGAGGCGCTTCCAGACGCGCCAAATC 715
 180 LeuLeuHisAlaArgGlnAspSerThrLyAspValAlaPheLeuIleThrAspGlyTyrSer 199
 716 CTTCGTCACTCTAGAGAAACTCCACCAAGTCAATATTTCTCATCAACGACGCTATATCC 775
 200 AsnGlyGlyAspProArgProGlnIleAlaAlaSerLeuArgAspSerGlyValGlnIlePhe 219
 776 AATGGCGGAGACCCCAAGACTATTTGACAGATCCCTTGGAGATTCGAGATGGAGATCTTC 835
 220 ThrPheGlnIleTrpGlnIleAsnIleArgGlnLeuLeuAspMetAlaSerThrProLys 239

836 ACCTTCGGGATTTGGACGGGAGATATCCGGGAACTGAATGACTATGCTTCCACCCCGAAG 895
 240 GlnGlnIleScyVTrLeuLeuHisSerPheGlnGlnIlePheGlnAlaLeuAlaArgArgAla 259
 896 GAAAGAACATGTGTACCTGCTCCACAGTGTGAAGAAATTTAGAGCTTTAGCTCGCAGGGCCG 955
 260 LeuHisGlnLeuLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
 956 TGGCATTAAGATTAACCTTCTGGAGAGTTTATTCAGAGAGATATAGCCCATGCTCTTAT 1015
 280 LeuCyAspArgGlnGlyLyAspCyCyAspArgMetGlySerCyLyCyAspGlyThrHis 299
 1016 CTCTGTAGCGTGGGAAAGACTGCTGTGACAGAAATGGCACTGGAAATGGGACACAC 1075
 300 ThrGlyHisPheGlnGlyIleCyGlnGlyGlyTyrTyrGlyGlyLeuGlnIleProGln 319
 1076 ACGGTCATTAATTCATCTGTGACAGAGGCTATTAACGGGAAAGGCTCGACAGATGAG 1135
 320 CysThrAlaCysProSerGlyThrTyrLyAspProGlnGlySerProGlnIleIleSerSer 339
 1136 TGCACACTTGCCTCATAGAGGACATATTAAGCGGAGCTTCTCCAGAGAGATCAAGACCC 1195
 340 CysIleProCysProAspGlnAsnHisThrSerProProGlySerThrSerProGlnAsp 359
 1196 TGCATCCCATGTCTGAAGTAAAGCACACCTCCCACTGGAAGCACTTCCCTGAAGAC 1255
 360 CysValCysArgGlnGlyTyrArgAlaSerGlyGlnThrCysGlnLeuValHisCysPro 379
 1256 TGCCTGTGCCCAAGAGGATACACAGATCTGGCAGACCTTGAGAGTGTCCACGTGCT 1315
 380 AlaLeuLysProProGlnAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
 1316 GCCCTGAAGCTCTGAAGAAATGAGTTTATTAACAAACCTTGCAAAATCACTTCAAT 1375
 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeu 419
 1376 GCGGCTGTGGGTCAGATGCGCGCGCTTAACTGTGAGGAGACAGATCCATTTG 1435
 420 CysLeuProAsnGlyLeuTrpSerGlySerGlySerTyrCysArgValArgThrCysPro 439
 1436 TGTCAACCAATGTTGTGTGTCTGTGACAGAAAGCTTCTCAGAGTGAAGACGTGCCCC 1495
 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGlnMetLeuTyrLys 459
 1496 CACTTCCAGACGCCAAACAGGCCACATCACTGCTCACTGCGGAAATGCTCTCAAC 1555
 460 ThrThrCysLeuValAlaCysAspGlnGlyTyrArgLeuGlnGlySerAspLysLeuThr 479
 1556 ACCGTGTGTTGGTTACTGTCAAGTAAAGATTAAGATTAAGAGGACAGCATAGGCTTACC 1615
 480 CysGlnGlnAsnSerGlnTrpAspGlyProGlnProArgCysValGlnIleCysSer 499
 1616 TGTCAAGGAAATGCGCAGTGGAGTGGCCAGAGCCCGGTGTGTGAAGACGCATGTGACC 1675
 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnGlyLyGlnIleProAla 519
 1676 ACCTTCCAGAGGCCCAAGAGGCGTATCTTCAACCACTGCGGAGAGAGAGCCGCC 1735
 520 LysPheGlnThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
 1736 AGGCTGTGAGTGAACCTGTCACTGAAGTGCAGGAGGATCAATTTATCCGGGGCTCAGA 1795
 540 GlnMetLeuArgCysThrThrSerGlyLySerThrAsnValGlyValGlnAlaValCys 559
 1796 GAA---GTGAGTGTGCACATCTGGGAAAGTGAAGTGCCAAAGTTCACACAGCTGTGTC 1852
 560 LysAspValGlnAlaProGlnIleAsnCysProLysAspIleGlnAlaLysThrLeuGln 579
 1853 AAGATGTGAGGCTTCAAAATCACTGTCTCAATATGACATTTAGAGCAAAAGCTGGGAG 1912
 580 GlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGln 599
 1913 CAGCAGGACTCTGTATATGCCACTGGCAAGTCCCAACGCTAAAGAACACTGTGTGAA 1972

QY 600 LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
 DB 1973 MAGSTCTCATGTCACCTCCACCAAGCTTTACCCCACTTACTCTTCCCAATTGGAGAC 2032
 QY 620 ValAlaIleValIleThrAlaThrAspLeuSerGlyValSnglnAlaSerCysIlePheHis 639
 DB 2033 GTGGCCATCACCCTACCGGCAACCGACTATCCGATACCAAGCCAGCTGCATCTTCTAC 2092
 QY 640 IleValValIleAspAlaGluProProValIleAspTyrCysArgSerProProVal 659
 DB 2093 ATTAAAGTCATTGATGATGAGACCGCTGTCATGATGATGGCGATCTCCACCTCCAAATC 2152
 QY 660 GlnValSerGlyLeuValHisAlaAlaSerTyrAspGlyLeuProGlnPheSerAspAsnSer 679
 DB 2153 CAGCTCTGAGAGAGAGAGACCCCTGCACTGGGAGTGAAGCTGATCTCCAGCAACTCC 2212
 QY 680 GlValAlaGluLeuValIleThrArgSerHisThrGlnGlyValAspLeuPheProGlnGly 699
 DB 2213 GGGGCTGAATTGGTCATTACACAGCACTCACACAAAGCGACATGTTTCTCTCATGGGGA 2272
 QY 700 ThrIleValGlnIleThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
 DB 2273 ACGGTGTGTGTACACGCACTGACCCCTCAGGCAACAGACCTGTGACATCCAC 2332
 QY 720 IleValIleLeuGlySerProCysGlyIleProPheThrProValSnglnAspPheIle 739
 DB 2333 ATTGTCTATTAAGAGTTCTCCCTGTAGAGTCCCTTCACTCCCTTAAACGGGACTTATC 2392
 QY 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGlnGlyTyrAspPhe 759
 DB 2393 TGTGCCAGGATGCTGAGATTAACTGATGACCTGACGCAAGGAGGCTATGATTC 2452
 QY 760 ThrGlnGlySerThrAspIleTyrTyrCysAlaTyrGlnAspGlyValIlePheIleProThr 779
 DB 2453 ACGAAGGGGTACCTAGAGAGTACTGCTTGTGAAGATGATATCTGAGACCA 2512
 QY 780 TyrThrThrGlyLeuThrProAspCysAlaIleValArgPheAlaAsnHisGlyPheIleSer 799
 DB 2513 TACTCTACAGAAATGGCCAGCTGTCTATAAAAGTTTTCACACATGTTTCAAGTCC 2572
 QY 800 PheGlnMetPheTyrLeuValAlaArgCysAspAspThrAspLeuMetIleValPheSer 819
 DB 2573 TTGAATGCTATACAAACACCTCGCTGTGATGACATGATCTGTTAAGAAAGTTTCT 2632
 QY 820 GluAlaPheGlnIleThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
 DB 2633 GCAGCATTTGAGACTACCTCGGAGAACATGCTCCCTTTGTATGATGATGATGAC 2692
 QY 840 IleAspCysArgLeuGlnGluAsnLeuThrLysIleTyrCysLeuGlnIleTyrAsnTyrAsp 859
 DB 2693 ATTGACTGCAGCTGAGAGAC---CTGACCAAAAATATCTGCATCGATTAATTACAC 2749
 QY 860 TyrGlnAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAsnArgLeuAspTyr 879
 DB 2750 TATGAATAATGGCTTGTGATTTGACACGAGAGCTGGGTGACGCAACAGCTGGATAT 2809
 QY 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
 DB 2810 TCTTACGATACCTTCGATGTTGTACAGAAACACCCACCATGTGGGAGAGGACGA 2869
 QY 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLeuIlePheAsn 919
 DB 2870 TCTTCAAGGATTAAGAAGACATGCTCCATGTCTGACCCCAAAATTCAGCTTAATTTTAC 2929
 QY 920 IleThrAlaSerValProLeuProAspGlyLysAsnAspThrLeuGlnIleTyrGlnAsnGln 939
 DB 2930 ATCACAAGCTAGCTGCTGACCTCCAGAGAAAGAAACATATCCCTGTGAATGGAGAAACAG 2989
 QY 940 GlnArgLeuLeuGlnIleThrLeuGlnIleThrAlaAsnLeuLysArgThrLeuAsnLys 959
 DB 2990 CAGCGACTACTTAAGACATTGAAACATCAACATGCTGTAAAGACACTTGAATTA 3049

QY 960 AspProMetTyrSerPheGlnLeuAlaSerGlnIleLeuIleAlaAspSerAsnSerLeu 979
 DB 3050 GAGCCCATGATATCTTCCAGCTGCTCGAAGAACATGTGGCTGACAGCAATTCCTTC 3109
 QY 980 GluThrLysAlaIleSerProPheCysArgProGlySerValLeuArgGlyLysArgMetCys 999
 DB 3110 GAAACAGAAAGGCTTTCTCTCTGACACAGGCTGTGTGTGAGGGGGCGGATGT 3169
 QY 1000 ValAsnCysProLeuGlyThrTyrTyrAsnLeuGlnIlePheThrCysGlnSerCysArg 1019
 DB 3170 GTCACTGCCCCCTGGGAACCTTACTCTCTGAGACATTCACCTGTAAAGCTGCTC 3229
 QY 1020 IleGlySerTyrGlnAspGlnGlnGlnIleGlnLeuGlnCysLysLeuCysProSerGlyMet 1039
 DB 3230 ATGGATCTCATCAAGATGAAAGAGGACGCTGGAATGCAAGCTGTGCTCCCAAGACT 3289
 QY 1040 TyrThrGlnIleHisSerArgAsnIleSerAspCysValAlaGlnCysLysGlnGly 1059
 DB 3290 CAGCGGAATACCTCAATTCAGAGACGCTGTGAATGCAAACTCACTGTAAAGAGG 3349
 QY 1060 ThrTyrSerTyrSerGlyLeuGlnIleThrCysGlnSerCysProLeuGlyThrTyrGlnPro 1079
 DB 3350 ACCTACTCTCCAGTGGCTGGAGACTGCGAATGCTGCTGCTGCTGCTTATCAACCG 3409
 QY 1080 LysPheGlySerArgSerCysLeuSerCysProGlnAsnThrSerThrValLysArgGly 1099
 DB 3410 GAATTGGATCCCGAGCTGCTCTCTATGCCAGAAACCAACCAACGCTGAAAGAGGA 3469
 QY 1100 AlaValAsnIleSerAlaCysGlyValIleProCysProGlnGlyLysPheSerArgSerGly 1119
 DB 3470 GCGGTGACATCTCTGTGTGTGAGTGCCTGCGCCAGTAAAGAAATCTCCGTTGAGG 3529
 QY 1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAlaAlaGlyLysAlaPhe 1139
 DB 3530 CTAAACCTCTGACTACCTTCCCTTCCAGATTAATACCAACCAATGAGGAGAGTCTTC 3589
 QY 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
 DB 3590 TGCCTGCTGTGCTCTTATATGAACTACCAACATCACTGCGCCACGCTCATCAACAG 3649
 QY 1160 CysSerSerPheSerSerThrPheSerAlaAlaGlnGlnIleValIleProProAlaSer 1179
 DB 3650 TGTCTCAAGTTTACCTTACTTCTCAGCAGCAAGAAAGACATAGTCCCTGTGGCC 3709
 QY 1180 LeuGlyHisIleLysLysArgGlyIleGlyIleSerSerGlnValPheIleGlyCysPhePhe 1199
 DB 3710 CTGACATTCCTCCAGAAAGATGAGAGACAGCTGAGGCTTTCAGATGCTTCTTA 3769
 QY 1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyLysArgIleTyrValCysLeuCys 1219
 DB 3770 AACCCCTGCAACAAGTGAACCTGCAACAGCTTGGGCGTGTATGTCTGTCTGTC 3829
 QY 1220 ProLeuGlyTyrThrGlyLeuLysCysGlnIleThrAspIleAspGlyCysSerProLeuPro 1239
 DB 3830 CACCTGATACACAGCTTAAGATGTAAGATTAATTAATGATGATGATGATGATGATGAT 3889
 QY 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyIlePheIleCysGlnCysProSer 1259
 DB 3890 TGCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3949
 QY 1260 GlyTyrThrGlyGlnIleArgCysGlnGlnAsnIleAsnGlyCysSerSerSerProCysLeu 1279
 DB 3950 GGTATTCAGATCAAAATATGTAAGAAATATTAATATGATATGATATGATATGATATGAT 4009
 QY 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
 DB 4010 AATTAAGAAACCTGCACTGAGGCTTGGAGACTACCGCTTACTGTGTAAAGATAC 4069
 QY 1300 ValGlyLeuHisCysGlnIleThrGlnValAlaSnglnCysGlnSerAsnProCysLeuAsnAsn 1319
 DB 4070 ATGGGTGTGACTGTGAACAAGCTCAATGAATGCAAGTCAACCCCTGCTTAACAAC 4129
 QY 1320 AlaValCysGlnAspGlnValGlyIlePheLeuCysLysCysProProGlyPheLeuGly 1339

Db 4130 GCGGTTGTAAGACCAAGTGGGGGCTTCGTCGCAATGCCCCAGATTTTGGCT 4189
 QY 1340 ThrArgCysGlyIleValAsnValAspGluCysLeuSerGlnProCysIleAsnGlyAlaThr 1359
 Db 4190 ACTCGGTGTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4249
 QY 1360 CysIleAsnGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
 Db 4250 TGTAAAGATGCTGCACACAGCTTCAGGTGTCATGTCACAGCAGCTTCACAGGAGACAC 4309
 QY 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
 Db 4310 TGTGAACGTAACATCAACAGATGTCAGTCCAAACCGGTAGAACACAGGACCTGGTGTG 4359
 QY 1400 AspGluLeuAsnSerIleSerCysIleCysGlnProGlyPheSerGlyIleValCysGlu 1419
 Db 4370 GATGAACCTAACTCATACAGTTGTAATGTCAGCCAGATTTTCAAGCCACAGGTGATG 4429
 QY 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleValCysVal 1439
 Db 4430 ACAGAACAGCTTCGCGTTTAACTGATTTGAAGTTCTGGCATCTACAGGTAACGTC 4489
 QY 1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetIleSer 1459
 Db 4490 CTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4549
 QY 1460 SerAspAspMetLeuValIleThrProIleSerValAlaValAspAsnGlySerAspAsn 1479
 Db 4550 TCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4609
 QY 1480 ThrLeuLeuLeuThrAspIleValIleValIleValIleValIleValIleValIle 1499
 Db 4610 ACCTTCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4669
 QY 1500 ThrAsnCysProSerValAlaAspAspGlyArgTrpHisIleIleAlaIleThrTrpSer 1519
 Db 4670 ACCAACGCGCCCTCCGTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4729
 QY 1520 AlaAsnGlyIleTrpIleValIleValIleValIleValIleValIleValIleValIle 1539
 Db 4730 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4789
 QY 1540 SerValGlyLeuProIleProGlyIleValIleValIleValIleValIleValIleVal 1559
 Db 4790 TCCATTGGCAAGCCATACCTGATGATGATGATGATGATGATGATGATGATGATGAT 4849
 QY 1560 LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579
 Db 4850 AAAG 4909
 QY 1580 TrpAspIleValIleLeuSerProGlnIleValIleSerLeuAlaThrSerCysProGlu 1599
 Db 4910 TGGGATCATATCTCTGTCACACAGATGATGATGATGATGATGATGATGATGATGATGAT 4969
 QY 1600 LeuSerIleGlyAsnValIleValIleValIleValIleValIleValIleValIleVal 1619
 Db 4970 CTGAGTCGAGGAAACGTTAGCATGCGCATTTCTGTCGAGGAAACACGCGGAAAGTGT 5029
 QY 1620 LysIleAsnSerIleSerIlePheCysSerAspCysProArgLeuGlyIleValIleVal 1639
 Db 5030 AAGGTGATTCGACGACATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 5089
 QY 1640 HisLeuArgThrAlaSerGluAspLeuValProGlySerIleValIleValIleValIleVal 1659
 Db 5090 CACTGAGACCTCATAGGAATCGAAACGAGCTCCAAAGTCACTGTTCTGATGAT 5149
 QY 1660 ProGlyPheGlnLeuValIleValIleValIleValIleValIleValIleValIleValIle 1679
 Db 5150 CCGGCGCTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5209
 QY 1680 GlnProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGly 1699

Db 5210 CAACCACTCCCACTGTGAACGATTCGCTGTGGCTCCCGCTTGGAGAAATGCG 5269
 QY 1700 PheHisSerAlaAspAspPheValAlaGlySerThrValIleThrValIleCysAsnAsnGly 1719
 Db 5270 TTCTACTCAGCCGAGAGACTTCATGCGGCGAGCAGCGGATCACTATCACTGACACAGTGGC 5329
 QY 1720 TyrTrpLeuLeuGlyIleAspSerThrMetPheCysThrAspAsnGlySerTrpAsnGlyVal 1739
 Db 5330 TACTACTGCTGGGTGATTCGGAATGTTCTGACAGCAACAGGAGCTGGAACGCGATTT 5389
 QY 1740 SerProSerCysLeuAspValAspGlyCysAlaValAlaGlySerAspCysSerGlyHisAla 1759
 Db 5390 TCACATCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5449
 QY 1760 SerCysLeuAsnValAspGlySerTrpIleCysSerCysValProProIleThrGlyAsp 1779
 Db 5450 TCTGCGCTGAACACCAACAGATTCCTAGATGATGATGATGATGATGATGATGATGATGAT 5509
 QY 1780 GlyIleAsnGlyAlaGluProIleValCysIleValAlaProGlyAsnProGlyIleVal 1799
 Db 5510 GGGAAAACTGTGAGAACTGTAAATGTAAAGCTCCAGAAAAATCCAGAAATGGCCGC 5569
 QY 1800 SerSerGlyGluIleValIleValIleValIleValIleValIleValIleValIleValIle 1819
 Db 5570 TCTTCTGCGAGATTTACACCGTGGTACTGATGATGATGATGATGATGATGATGATGATGAT 5629
 QY 1820 GlnLeuMetGlyValIleThrValIleThrCysLeuGluSerGlyIleValIleValIle 1839
 Db 5630 GAGCTGGTGGAGTGAAGACATCAAGTGTGAGACTGCGAGTGGATGCGCTCAGG 5689
 QY 1840 ProTrpCys 1842
 Db 5690 CCGTCTGT 5698
 RESULT 7
 LOCUS AX375163 11230 bp DNA linear PAT 01-MAR-2002
 DEFINITION Sequence 3 from Patent WO0210388.
 ACCESSION AX375163
 VERSION AX375163.1 GI:19169917
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCES
 1. Welcher A.A. and Billie G.S.
 C3b/c4b complement receptor-like molecules and uses thereof
 JOURNAL Patent: WO 0210388-A 3 07-FEB-2002;
 Amgen Inc. (US)
 FEATURES
 source location/Qualifiers
 1..11230
 /organism="Mus musculus"
 /mol_type="unasigned DNA"
 /db_xref="taxon:10090"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 11230
 Score: 8652.50 Matches: 1532
 Percent Similarity: 90.56% Conservative: 137
 Best Local Similarity: 83.13% Mismatches: 171
 Query Match: 84.67% Indels: 3
 DB: Gaps: 3
 US-09-977-053-6 (1-1842) x AX375163 (1-11230)
 QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValIleSerGlyTrpAlaThr 20
 Db 176 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 235
 QY 21 PheGlnIleMetSerProSerArgAspAspPheArgLeuPheProGluThrAlaPro 40

Db 226 TTCACGCCGGGCGCCCTTCGCTCAACTTCAGCTTCGCCCTGTTCCCGAGGCCCTCCG 295
 Qy 41 GtYAlaProGlySerIleProAlaProAlaProGly---AspGluAlaIleGlySer 59
 Db 226 GGGGCTCTGGGAGAGACTGGCGGATACCTCCGCTCCAGTGAAGAGAGAGGAGAGGAGAG 355
 Qy 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSer 79
 Db 356 AAGGTGAACGCGCTGGGCGCGGCTTCCGAGACGCGGTGGCGGACTGCGGAGCTCAGC 415
 Qy 80 GluArgLeuGluLeuValPheLeuValAlaPheSerSerSerValGlyGluValAlaPhe 99
 Db 416 GCGAGCTGAGCTGCTCTTCCTGGTGGAGAGTCTCCAGCTGAGGACCAACTTC 475
 Qy 100 AcgSerGluLeuMetPheValArgGlyLeuLeuSerAspPheProValProThrAla 119
 Db 476 CTCACAGAGCTCAAGTTCCTGGCCAGAGCTCTCCAGCTTCGCCGTGGTCCAGGCC 535
 Qy 120 ThrArgValAlaIleValThrPheSerSerLeuAsnTrpValValProArgValAspTrp 139
 Db 536 AGCGGTGGCCATCGTCACTTCATCCAGAACACAGTGGTGGCGCGCTGATTC 595
 Qy 140 IleSerThrArgArgAlaArgGlnHisGlyGlyAlaLeuLeuGlnGlnIleProAla 159
 Db 596 ATCTCCACAGCGCGCGGACCAACACAGTGGCGCTGCTCCAGCGCGAGATCCGCGC 655
 Qy 160 IleSerTrpArgGlyGlyGlyThrTrpTrpArgValAlaPheGlnGlnAlaGlnIle 179
 Db 656 ATCACTACCGCGGTGGTGGCACTATCCAGAGGCGCTTCCAGAGCGCGGCAATC 715
 Qy 180 LeuLeuHisAlaArgGluAsnSerThrTrpValValPheLeuIleThrAspGlyTrpSer 199
 Db 716 CTTCGCTCACTAGAGAAATCACCMAAGTCATTTCTCATCACCGACTATTC 775
 Qy 200 AsnGlyGlyAspProArgProIleAlaIleSerLeuArgAspSerGlyValGlnIlePhe 219
 Db 776 AATGGGAGAGCCCAACACTATTCAGACATCGCTTCGAGATTCGAGGAGATCTTC 835
 Qy 220 ThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnMetAlaSerThrProLys 239
 Db 836 ACCTTCGGGATTTGGAGGGGAAATCCGGGAACTGAATGACATGGCTTCCACCCGAG 895
 Qy 240 GlnGlnHisCysTrpLeuLeuHisSerPheGlnGlnPheGlnAlaLeuAlaArgAla 259
 Db 896 GAAGAACTGTATACCTGCTCCACAGTTTGAAGATTTGAGGCTTACCTCCACGCGC 955
 Qy 260 LeuHisGluAspLeuProSerGlySerPheIleGluAspMetValHisCysSerTrp 279
 Db 956 TTGCATGAAGATCTACTCTTCGAGGATTTATCCAGAGATGGCCCACTCTCTAT 1015
 Qy 280 LeuCysAspGluGlyLeuAspCysCysAspArgMetGlySerCysGlyCysGlyThrHis 299
 Db 1016 CTCTGTGAAGCTGGGAAAGATCTCTGACAGATGGCCAGCTGCAAAATGTGGAGACAC 1075
 Qy 300 ThrGlyHisPheGlnCysIleCysGlyGlyGlyTrpTrpGlyGlyGlyLeuGlnTrpGlu 319
 Db 1076 ACGGGTCAATTGAAATGATCTGTGAGAGGGCTATTACGAGAAAGGTCTGACGATGAG 1135
 Qy 320 CysThrAlaCysProSerGlyThrTrpLysProGluGlySerProGlyGlyIleSerSer 339
 Db 1136 TCACACAGCTTCCTCACTAGACCACTCCCACTGGAGCACTTCCCTGGAAGAC 1255
 Qy 340 CysIleProCysProAspGluAsnHisIleHisSerProProGlySerThrSerProLys 359
 Db 1196 TGCATTCATGCTCTACCTAGACCACTCCCACTGGAGCACTTCCCTGGAAGAC 1255
 Qy 360 CysValCysArgGluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
 Db 1256 TCGGTGTGCGAGAGGATACACAGATCTGCGACACCTGAGGTGTGCTCACTGTCC 1315
 Qy 380 AlaLeuLysProProGluGlnGlyTrpPheIleGluAsnThrCysAsnAsnHisPheAsn 399
 Db 1316 GCCGTGAAGCTCTGAAATGATTTTATACAAAACCTTGCAAAAACCTTCAAT 1375

Qy 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeu 419
 Db 1376 GCGGCTGTGGGATCCGATGTCGCGGCTTTGACTGTGGAGAGACATCTCATTTG 1435
 Qy 420 CysLeuProAsnGlyLeuTrpSerGlySerGlnSerTrpCysArgValAlaGluTrpCysPro 439
 Db 1436 TGTCAACCAATGGTTGGTCTGGAGCAAAAGCTTGTGCAGAGTGAACGTCGCC 1495
 Qy 440 HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTrpLys 459
 Db 1496 CACTTCGACGCGCAACACGCGCACTCAGCTGCTCACTCGGAAATGCTTACAC 1555
 Qy 460 ThrThrCysLeuValAlaCysAspGluGlyTrpArgLeuGlnGlySerAspLysLeuThr 479
 Db 1556 ACCCTGTGTTGGTACTCGCAATGAGATACAGATTGAGAGGAGCACTAGGCTTACC 1615
 Qy 480 CysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSer 499
 Db 1616 TGTCAAGAAATGCCAGTGGAGTGGCCAGAACCCCGGTGTAGAACCCATTTGCC 1675
 Qy 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
 Db 1676 ACTTTCAGAGCCCAAGCGCTCATCTTCCACACGAGTGGCGCAAGCGCGCC 1735
 Qy 520 ThrPheGlyThrIleCysTrpValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
 Db 1736 AGGCTGGAGATGACTGTCACTGATGCTGCGGCAAGGATACATTTATCCGGGCTCAGA 1795
 Qy 540 GlnMetLeuArgCysThrTrpSerGlyLysTrpAsnValGlyValGlnAlaValCys 559
 Db 1796 GAA--GTAGAGTGTCCACATTTGGAGTGAAGTGAAGCCMAATTCAGACAGCTGTGTGC 1852
 Qy 560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGln 579
 Db 1853 AAGATGTGAGGCTCCAAATCAGCTGTCCAAATGACATGAGCAAAAGATGGGAG 1912
 Qy 580 GlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlu 599
 Db 1913 CAGCAGACTCTCTAATGCCACCTGCAAGTCCCAACAGCTMAAGCACTGTGGAA 1972
 Qy 600 LysValSerValHisValHisProAlaPheThrProProTrpLeuPheProIleGlyAsp 619
 Db 1973 AAGGTTCAGTCCACCTCCACCGCTTTACCCACCTTACCTTCCAAATGGAGAC 2032
 Qy 620 ValAlaIleValTrpAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
 Db 2033 GTGGCATCACCTACAGCGCAACCGATCATCGGTACCAACGAGCTGCACTTCTAC 2092
 Qy 640 IleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProProProVal 659
 Db 2093 ATTAAGTCAATGATGTGAACCGCTGATGATGTTGGTCCGATCTCCACTCCAAATC 2152
 Qy 660 GlnValSerGluLysValHisAlaIleSerTrpAspGluProGlnPheSerAspAsnSer 679
 Db 2153 CAGGTGTGAAGAGAGGACCTCGAAGCTGGAGTGAAGCTCAAGTTCTGACACACTCC 2212
 Qy 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
 Db 2213 GGGGCTGAATGTCTATTCACAGACTCACACAAAGGAGCAATGTTCTCTCATGGGAA 2272
 Qy 700 ThrIleValGlnTrpThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
 Db 2273 ACGGTGTGTGTACACAGCACTGACCCCTCAGGCAACAAAGGACCTGTGACATTCAC 2332
 Qy 720 IleValIleLysGlySerProCysGlnIleProPheThrProValAsnGlyAspPheIle 739
 Db 2333 ATGTGATTAAGATCTCCCTGTGAGTCCCTTCAACCCCTTAACCGGAGCTTATTC 2392
 Qy 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGlnGlyTrpAspPhe 759
 Db 2393 TGTGCCAGAGATGTCCTGAGATTACTGATGCTGAGCTGCAAGAGGAGCTATGATTTTC 2452

```

QY 760 ThrGluGlySerThrAspLeuYrYrYrCysAlaIleTyrGluAspGlyValIleProLysProThr 779
DB 2453 ACGAGAGGGGTCACTGAGAGAGTACTGCTTTTGAAAGATGCTATCTGAGACCA 2512
QY 780 TyrThrThrGluTyrProAspCysAlaIleYsAspPheAlaAsnHisGlyPheLysSer 799
DB 2513 TACTCTACAGATGCGCAGACTGTCTATTAACGTTTTCGCAACCATGCTTCAATCC 2572
QY 800 PheGluMetPheTyrIleYsAlaAlaIleArgCysAspAspThrAspLeuMetLysLysPheSer 819
DB 2573 TTTGAATGCTATACAAACCACTCGCTGATGACATGATCTGTTAAGAGTTTCT 2632
QY 820 GluAlaPheGluThrThrIleGluGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
DB 2633 GCAGCTTTGAGACTGACCTCGGAGAACATGCTCCGCTTTGTAAACGATGCGATGAC 2692
QY 840 IleAspCysArgLeuGluGluAsnLeuThrLysIleYrYrCysLeuGluTyrAsnTyrAsp 859
DB 2693 ATTGACTGCGAGCTGGAGAC---CTGACCAAAAATATCTGATCGATATATTAACAC 2749
QY 860 TyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAlaAsnArgLeuAspTyr 879
DB 2750 TATGAATAATGGCTTTGGCAATTGACCAAGAGCTGGGTGACAGCAACAGCTTGATTA 2809
QY 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
DB 2810 TCTTACGATCACTTCGATGATGTGACAGAAACACCCAGATGCGAGAGGACGAGA 2869
QY 900 SerSerArgIleIleYsAspSerAlaProLeuSerAspTyrIleYsIleLysLeuIlePheAsn 919
DB 2870 TCGTCAAGATTAAGAAACATGCTCCATTTCTGACCCCAAAATTCAGCTAATTTTAC 2929
QY 920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrIleGluTyrGluAsnGln 939
DB 2930 ATCAGAGCTAGCGGTGACCTCCAGAGAAAGAAACATACCTTGAAATGGAGATCAG 2989
QY 940 GlnArgLeuLeuGlnThrIleGluGluThrIleThrAsnLysLeuLysArgThrIleAsnLys 959
DB 2990 CAGCGACTCATTAAGACATTAAGAAACATACCATGCTGAGAAACACCTTGAAATTA 3049
QY 960 AspProMetTyrSerPheGluIleLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
DB 3050 GAGCCCATGATCTTTCAGCTGCTCCCTCGAACAAGTGGGTGACAGCATTTCCCTC 3109
QY 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgIleArgMetCys 999
DB 3110 GAAGAAGAAAGGCTTTCTCTCTGCAACAGGCTGTGCTGAGGGGGCGCATGT 3169
QY 1000 ValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArg 1019
DB 3170 GTCAACTGCCCCCTGGAGAACCTTACTCTCTGAGACATTCACCTGGAAGCTGCTC 3229
QY 1020 IleGlySerTyrGluAspGluGluGluIleGluLeuGluCysLysLeuCysProSerGlyMet 1039
DB 3230 ATGGAGTCTTACAGATTAAGAGAGGAGCTGAGATGACAGCTCTGCTCCCAAGACT 3289
QY 1040 TyrThrGluTyrIleIleIleSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
DB 3290 CAGCGGATATACCTCATTAAGAAAGCTCTGAAATGCAAAAGCTCAGTGTAAACAAGGC 3349
QY 1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
DB 3350 ACTTACTCTTCAGTGGGCTGGAGAACCTGCGAATGCTGCGTGGGTATCTTAACAACG 3409
QY 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
DB 3410 GAATTTGATCCCGAGCTGCTCTATGCGCAAAACACACACAGCGTAAAGAGCA 3469
QY 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
DB 3470 GCCGTGACATCTCTGCTTGAGAGTCCCTGCGCAGATGAGAGAAATCTCCCGTTTCGG 3529
QY 1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
DB 3530 CTAAACCTCTGTAACCTTGTGCTCGAGACTATTAACCAACCATGACGGAATGCTTC 3589
QY 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
DB 3590 TGCTGCTGTGCTCTTTATGAACTACAAACCATCACTGGCGCAGCTCATACAGAC 3649
QY 1160 CysSerSerPheSerSerThrPheSerAlaAlaGluGluSerValIleProProAlaSer 1179
DB 3650 TGCTCAAGTTTACCTTACTCTCTGAGCAGCAAGAAAGAAATGAGCTCCCTGCTGCC 3709
QY 1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValIlePheHisGluCysPhePhe 1199
DB 3710 CTGAGACTTCCAGAAACAGATACAGATGACAGCTGAGCTTCTTCAAGATGCTTCTTA 3769
QY 1200 AsnProCysHisAsnSerGlyThrCysGlnGluLeuGlyArgGlyTyrValCysLeuCys 1219
DB 3770 AACCTCGCAACAAAGTGAACCTGCAACAGCTTGGGGGTGTTATGTCTGTCTGC 3829
QY 1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
DB 3830 CCACCTGATACAGAGCTTAAGTGTAAACATATTTGATGATGACAGCTCTCTGCT 3889
QY 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
DB 3890 TGCTCATGTGGAAATTTGTAAGACCAAGTTGGGATTCACGTGCAATGTTCTATG 3949
QY 1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu 1279
DB 3950 GGCTATTCAGTCAATATGTGAAGAAATATTAATGATGATATCTCAGCCCTTGCTTA 4009
QY 1280 AsnLysGlyLysCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
DB 4010 AATTAAGAAACCTGACCTGACGCTTGCAAGCTTACCGCTGATGTAAGAAATAC 4069
QY 1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
DB 4070 ATGGGTGCTGCTGTAAGAAACAGCTCATATATGCTGATGACCTCTGTTAAACAC 4129
QY 1320 AlaValCysGluAspGluValGlyGlyPheLeuCysLysCysProProGlyPheLeuGly 1339
DB 4130 GCACTTTGTAAGCAAAAGTTGGGGGTTCGTGCAAAATCCCAACCGGATTTTGGGT 4189
QY 1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyValThr 1359
DB 4190 ACTGCTGTGAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4249
QY 1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
DB 4250 TGTAAAGATGTGCAACAGCTTCAAGTGTCAATGTCCAGCAGGCTTCAAGGACACAC 4309
QY 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
DB 4310 TGTAACTGAACATCAACAGAGTCAATGCAACCGGTGTAAGAACAGGCCACCTGTGTG 4369
QY 1400 AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyValArgCysGlu 1419
DB 4370 GATTAACCTAACTATCACTGATTAAGTCAAGCAGGATTTTCAAGGCAAGGTGTAG 4429
QY 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
DB 4430 ACGAAGACGCTTCGGGTTTACCTGAGATTTGAAGTTTCTGCAATCTACAGGATACCTC 4489
QY 1440 MetLeuAspGlyMetLeuProSerIleHisAlaLeuThrCysThrPheTyrMetLysSer 1459
DB 4490 CTGCTAATGATGAGTGCAGCAACCTCATGCTCAATCCGTAACCTGCGCATTCGATGAATCC 4549
QY 1460 SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn 1479
DB 4550 TCTGATGTCATCAACTACGGAGACGCCCATCTCATATGACCTTGAGAGTACAAAGAACAC 4609
QY 1480 ThrLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGlyLysIle 1499

```

Db 4610 ACCTTCTCTGACGATTAACAAGCGCTGGGTTCTTATGTAAGCAAGAAAGATC 4669
 Qy ThrAenCyProSerValAenAapGlyArgTPhIshE1a1a1eThrTrpTherSer 1519
 Db 1500 ThrAenCyProSerValAenAapGlyArgTPhIshE1a1a1eThrTrpTherSer 1519
 Db 4670 ACCAATGCTCCCTCCGTAATATATGCGATTTGGCATCATATTCGATCATATGCAAT 4729
 Qy AlaAenGlyLeuTrpValValTyrTleAapGlyLysLeuSerAapGlyValAagLysLeu 1539
 Db 4730 ATTGAGGAGCGCTGAGGAGGCTATATAGATGAGGAAATATATGACGCTGCTACTGCGCTC 4789
 Qy SerValGlyLeuProIleProGlyValValAalaValLeuGlyValGlnGlnLysPhe 1559
 Db 4790 TCCATATGCAAAAGCATATCTGCTGGCGGTGATTAAGTCTTGCGCAAGACAAAGCAAA 4849
 Qy LysGlyGlnGlyPheSerProAagLysPheValGlySerTleSerGlnLeuAenLeu 1579
 Db 4850 AAGGAGAGGGGTTCACCGGCTGAGTCTTTTGCGCTCCATAGCCAGCTCAACCTC 4909
 Qy TrpAapTyrValLeuSerProGlnGlnValLysSerLeuA1eThrSerCyProGlnGln 1599
 Db 4910 TGGAGCTATGCTCTGCTCCACAGCGAGTGAAGTTCGCGCAGCTCCTGCCAGAGAA 4969
 Qy LeuSerLysGlyAanValLeuA1eTrpProAapPheLeuSerGlyTleValGlyLysVal 1619
 Db 4970 CTGAGTCCGGGAAAGCTGTACATGCGCCGATTCCTGCTGGGAAATCAAGGGAAAGTG 5029
 Qy LysIleAapSerLysSerIlePheCySerAapCyProArgLeuGlyLysSerValPro 1639
 Db 5030 AAGGTGATTCACAGACAGATGCTCTGCTGATGCTGCTCTTGAAGAGATCCGCGCT 5089
 Qy HisLeuArgThrAlaSerGlyLysApleuLysProGlySerLysValAanLeuPheCyAap 1659
 Db 5090 CACCTGAGAGCGGATCAGAAATCGAAAGCCAGGCTCCAAAGTCAGTCTGCTGAT 5149
 Qy ProGlyPheGlnLeuValGlyAanProValGlnTyrCyLeuAanGlnGlnGlnTyrPhe 1679
 Db 5150 CCGGCTTCACATAGTGGGAAATCTGCGAGTATGCTGAAACCAAGGCGAGGAGCA 5209
 Qy GlnProAapProHisCyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 1699
 Db 5210 CACACATCTCCCACTGTGAAGCATTCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 5269
 Qy PheHisSerAlaAapAapPheTyrAlaGlySerThrValThrTyrGlnCyAanAanGlnGly 1719
 Db 5270 TTCTACTCAGCGAGGACTTCATGCGGCGAGCAGCAGGATCATGATGACACAGCTGCG 5329
 Qy TyrTyrLeuLeuGlyAapSerArgMetPheCyThrAapAanGlySerThrAanGlyVal 1739
 Db 5330 TACTACCTGCTGGGATTCCTGGAATGCTCTGCRCAAGACAGGAGCTGGAAAGCGCAT 5389
 Qy SerProSerCyLeuAapValAapGlnCyAalaValGlySerAapCySerGlnLysIleAala 1759
 Db 5390 TCACATCTCTGCTGATTCATGATGCTGAGTCTGCGCTGCGCTGATGATGACAGCGCC 5449
 Qy SerCyLeuAanValAapGlySerTyrTleCySerCyValProProTyrThrGlyAap 1779
 Db 5450 TCCTCTCCTGAACACACAGGATTCATGATGCTCTGCTGATGACACACAGGAGAT 5509
 Qy GlyLysAanCyAalaGlnProLysGlyValAapProGlyAanProGlnLysIleAala 1799
 Db 5510 GGGAAAACTGCGAGAACCTGTAATATTAAGCTCCGAAATTCAGAAATTAAGCGCGC 5569
 Qy SerSerGlyGlnLysTyrThrValGlyAalaValThrPheSerCyGlnGlnGlnGlnGln 1819
 Db 5570 TCTTCTGCGAGATTAATTCACCGTGGGTACGCAAGTCAATTTTCGATGACAAAGGAGC 5629
 Qy GlnLeuMetGlyValThrLysIleThrCyLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 1839
 Db 5630 GAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5689
 Qy ProTyrCy 1842
 Db 5690 CCGTCTCTG 5698

RESULT 8
 AX880905
 LOCUS
 DEFINITION
 AX880905
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Ota, T., Iwagai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
 Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
 Primers for synthesizing full-length cDNA and their use
 Patent: JP 1074617-A (1998) 07-PEB-2001;
 Research Association for Biotechnology (JP)
 Location/Qualifiers
 1..5124
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 119..4069
 /note="unassigned protein product"
 /codon_start=1
 /protein_id="CA891130.1"
 /db_xref="GI:40035642"
 /translation="MASTPKEHCHYLHSEFEALARAALHEDLPSSGFIODMVC
 SYLCBGRDCDPMSCCKGTHGHEHCICGKYVKGKQYBCTACPSGTYKREGSPG
 GISSCTCPDENHNTSPGSTSPBDCYCRGASGTCGLVHCPALKPENNGPIONT
 CNHFNPAAGVCHPFDLVSSILCLNGLMSGLSTCYRATCPHLNPKHGLISC
 STREMLKYTCVACDGYRLRGSDPLTCOGSOMGSPRCVCHSCSTFOMKEDVI
 SHNCGOKPAKGTICVYSCROGFLISGVKMLACTTSGKMYGVAAVCKDEAVDI
 NCPKIDAKTLEADDSANTWQIPTRAKMSSEKSVHVAHPATPPLPIGDAIYVT
 APTDSGNQSCIPHIKVTDARPPVTDKMSPPVYSEKVAHMAHPDSDMSGAL
 VTRSHQGLDPQGETIYQYATTPDSGNRRCTDIHYIKSGSCBPIPTPVNDICT
 PNTGVNCTLCTLSGDTFBSSTDKYCAVEBGVMPPTVITTEPDKAKAFANHGRS
 PMFVYAAACDITDLAKKFSBAPEITLGMVSPSCDADIDRLLEBNITKYLCLBYN
 YDENGFAIPGQMGANRLDYSYDFLOTVEBTASIGNASRIRKSPVLSYKIK
 LFNITASVPLPDERNDITLERNQGLDITLETITKRLTKLNDKDMYSFOLASBIL
 ADSNLGTTKASAPCRPSVLRGMCVNLGTYVLEHTGSCSGIGSYODRGOLG
 CGLCPSGMARTYTHSRNISTDCRAQCKGCTYSCSGLETSCSPIGTYOPRSGRSCLC
 PNTSTVKGAVNISACVPCBGRSRSGMLCHPCEPDYTOPAKAPCLACPYNG
 TTPFAGSRSTIECSSTPSSTPSAABSVSPASIGHKKHHSQVPHBEPFPCNHS
 GTCOOLARGVCLCELYTGLKCTEDIDBCPLCLANGCVDLVBEPICSCSGYTG
 ORCERNINCSPPCLNKGIQVGVAGYCTCYKGVGLHCEFEVNECSNPGANNAV
 CBDQVGRPCRCRPGRIETGRCKNDKDELSOPCKNGATCKDANSFRCSCAAGFTGS
 CELATNECSNFCRNQATCTVDRLNYSCKCQCPGSPGRCRTBSGTPNIDFRVSGIY
 YVNLVGMPLPSIALCTTFMFKSSDDNVTGPTSYAVDQSDNTLLTIDTRGATLYANG
 RKRTNCSVNDGRMHLAITWTSANGIKVYIDKLSDGAGLSVGLPIPGMP"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 5124
 Score: 7537.00 Matches: 1347
 Percent Similarity: 99.63% Conservative: 1
 Best Local Similarity: 99.56% Mismatches: 5
 Query Match: 73.75% Indels: 0
 DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x AX880905 (1-5124)

Qy 195 ThrAapGlyTyrSerAanGlyGlyAapProArgProIleAalaAalaSerLeuAaapSer 214
 Db 2 ACTGATGATATTCGAATGGGGAAGACCTTAAGCAATATGACAGCTGACGAGATTC 61
 Qy 215 GlyValGlnLysPheThrPheGlyTleTyrGlnGlnGlnGlnGlnGlnGlnGlnGln 234
 Db 62 GAGATGAGATCTTCATTTTGCAATATGCAAGGAAACATTCGAGAGCTGAAATGACATG 121
 Qy 235 AlaSerThrProGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 254

Db 122 GCTTCACCCCAAGAGAGCACTGTAACCTGCTACACAGTTTGAAGATTGAGGCT 181
 Qy 255 LeuAlaArgArgAlaLeuHiSegLuePLeuProSerGlySerPhe11eGlnAspAspMet 274
 Db 132 TTAGCTCCGCGGGCATTCATGATGATTAACCTTCGGGAGTTTATTCAAGATGATATG 241
 Qy 275 ValHisCysSerThrLeuCybAspGluGlyValAspCysCysAspArgMetGlySerCys 234
 Db 242 GTCCACTGCTCATATCTTGTGTGATGAAAGGCMAGACTGCTGTGATCCGAATGGGAACTGC 301
 Qy 225 LysCysGlyThrHisThrGlyHisPheGluCys11eCysGluLysGlyTyrTyrGlyLys 314
 Db 302 AATGTGGAGACACACAGGCCATTTTGAAGTCATCTGTGAAGAGGGATTTTCGGGAAA 351
 Qy 315 GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro 334
 Db 352 GGTCTGAGATTAGATGACACAGCTTGCCATCGGGACATACAAACTGMAAGCTCACCA 421
 Qy 335 GlyGly11eSerSerCys11eProCysProAspGluAsnHisThrSerProGlySer 354
 Db 422 GGAGGAATTCAGACATTCATTCATGCTGATGAAATACACACTCTCCACCTGAGAGC 481
 Qy 355 ThrSerProGluAspCysVal1CysArgGluGlyTyrArgAlaSerGlyGlnThrCysGlu 374
 Db 482 ACATCCCTGAAAGACTGTGTCTGCAAGAGAGATACAGGCATCTGGCAGACTGTGA 541
 Qy 375 LeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPhe11eGlnAsnThrCys 394
 Db 542 CTGTTCACATGCTCTCCCTGAAAGCTCCCGAAATAGTTACTTTTCAAAACACTGTC 601
 Qy 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
 Db 602 AACCAACCACTTCAATGCAAGCTGTGGGCTCCGATGTCACCTTGATTTATCTTGGGGA 651
 Qy 415 SerSer11e1eLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArg 434
 Db 652 AGCAGCATCATCTTATGTTTACCCCATAGTTTGTGTGCTTCGTTTAAAGACTTCTGAGA 721
 Qy 435 ValArgThrCysProHisLeuArgGlnProLysGlyHisGly11eSerCysSerThrArg 454
 Db 722 GTAAAGAACATGCTCATCTCCGCGCAGCGAAACATGGCCACATCATCTGTCTACAGG 781
 Qy 455 GlnMetLeuTyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly 474
 Db 782 GAATATGTTATATAGACCAATGTTGTGTGCTGTGATGAAAGGTACAGACTAGAGGC 841
 Qy 475 SerAspLysLeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysVal 494
 Db 842 AGTGATTAAGCTTACTGTGCAAGGAAACAGCCAGTGGAGTGGCCAGAACCCCGGTGTG 901
 Qy 495 GlnArgHisCysSerThrPheGlnMetProLysAspVal11e1eSerProHisAsnCys 514
 Db 902 GAGGCCCATCTGTTCACCTTTCAAGATGCCCAAGATGATCATCATATCCCCCAACATCTG 961
 Qy 515 GlyLysGlnProAlaLysPheGlyThr11eCysTyrValSerCysArgGlnGlyPhe11e 534
 Db 962 GGCAGACGACCGCCAAATTTGGGACGATCTGTATGTAAGTTGCGCGCAAGGTTTCATT 1021
 Qy 535 LeuSerGlyValLysGlnMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyVal 554
 Db 1022 TTATCTGAAGTCAAGAAATGCTGTGAGATTAACATCTTGGAATAATGGAATGTCCGAGTT 1081
 Qy 555 GlnAlaAlaValCysLysAspValGluAlaProGln11eAsnCysProLysAsp11eGlu 574
 Db 1082 CAGGACGCTGTGTGAAGACGTGAGGCTCCCTCAATCAACTGTCTTAAGGACATAGAG 1141
 Qy 575 AlaLysThrLeuGlnGlnLysPheAlaAsnValThrTrpGln11eProThrAlaLys 594
 Db 1142 GCTAAGACTCTGGAAACAGCAAAATTCCTGCATGTTAACCTGGCAGATTCACACAGCTAA 1201
 Qy 595 AspAsnSerGlyLysValSerValHisValHisProAlaPheThrProProTyrLeu 614
 Db 1202 GACAACTTGTGTGAAGAGGTGACGTCCAGCTTCACTCCAGCTTTCACCCCACTTACTT 1261

Qy 615 PhePro11eGlyAspValAla1eVal1TyrThrAlaThrAspLeuSerGlyAsnGlnAla 634
 Db 1262 TTCCCAATTGGAGATGTTCATGCTATACAGGCACTGACTATCCGGCAACAGGCTC 1321
 Qy 635 SerCys11ePheHis11eLysVal11eAspAlaGluProProVal11eAspTrpCysArg 654
 Db 1322 AGCTGCAATTTTCCATTCATAGGTTTATGATGCAAGACCACTGTGCATAGACTGTGCA 1381
 Qy 655 SerProProValGlnValSerGluLysValHisAlaAspTrpAspGluProGln 674
 Db 1382 TCTCCACTCCCGTCCAGGCTCGAGAGAGGTACATGCCCAAGCTGGATGACCTCAG 1441
 Qy 675 PheSerAspAsnSerGlyValGluLeuVal11eThrArgSerHisThrGlnGlyAspLeu 694
 Db 1442 TTCTCAGACAACTCAGGGGCTGAATTTGTGATTAACGGAAGTCAATACAGAGGACTT 1501
 Qy 695 PheProGluGlyGluThr11eValGlnTyrThrAlaThrAspProSerGlyAsnAsnArg 714
 Db 1502 TTCCCTCAGGGAGACTATAGTACGTATACGGCACTGACCTTACGGCAATACAGG 1561
 Qy 715 ThrCysAsp11eHis11eVal11eLysGlySerProCysGlu11eProPheThrProVal 734
 Db 1562 ACATGTGATATCATATTCATTAAGAGTTCTCCCTGTGAATTCATTCACACCTGTA 1621
 Qy 735 AsnGlyAspPhe11eCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeu 754
 Db 1622 AATGGGGAATTTATATGCACTTCAGATTAATACAGACTCAACTGATATTAACCTGTG 1681
 Qy 755 GlnGlyTyrAspPheThrGlnGlySerThrAspLysTyrTyrCysAlaTyrGluAspGly 774
 Db 1682 GAGGGCTATGATTTCAACAGAGGGTCTTACTGACAGATTAATTTGTCTTATGAAGAAGC 1741
 Qy 775 ValTrpLysProThrTyrThrThrGlnTrpProAspCysAlaLysLysArgPheAlaAsn 794
 Db 1742 GTCTGAAACCAACATATACATGATGCAATGCAAGTCACTGACCAAAACGTTTGCAAC 1801
 Qy 795 HisGlyPheLysSerPheGlnMetPheTyrLysAlaAlaArgCysAspAspThrAspLeu 814
 Db 1802 CACGGGTTCAAGTCTTGTGATGATTTCTCAAGCAAGCTGTTGTATGACAGACTGTG 1861
 Qy 815 MetLysLysPheSerGlnAlaPheGluThrThrLeuGlnLysMetValProSerPheCys 834
 Db 1862 ATGAGAGAGTTTCTGAGGATTTGAGACCACTCGGGAATAATGTTCCATCATTTTGT 1921
 Qy 835 SerAspAlaGluAsp11eAspCysArgLeuGluGluAsnLeuThrLysTyrCysLeu 854
 Db 1922 AGTGATGAGAGACATTTACTGAGACTGAGAGAGAACTGACACAAATAATGCTTA 1981
 Qy 855 GlnTyrAsnTyrAspTyrGluAsnGlyPheAla11eGlyProGlyGlyTyrGlyValAla 874
 Db 1982 GAATATATATATATATGAAATGCTGTGCAATTTGACAGAGTGTGGGGTGCAGCT 2041
 Qy 875 AsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSer 894
 Db 2042 AATAGCTGATTAACCTTACATGATGACTTCTGAGACTGTGCAAGAAACAGCCACAAGC 2101
 Qy 895 11eGlyAsnAlaLysSerSerArg11eLysArgSerAlaProLeuSerAspTyrLys11e 914
 Db 2102 ATCGGCAATCCCAAGTCTTCAGAGATTAAAGAGAGGCCCATTAATCTGACTTAATAATT 2161
 Qy 915 LysLeu11ePheAsn11eThrAlaSerValProLeuProAspGluLysAsnAspThrLeu 934
 Db 2162 AAGTTAAATTTTAACTCAACGCTAGTGTGCCATTTCCCATGAAAGAAATGATACCTT 2221
 Qy 935 GlnTrpGluAsnGlnGlnArgLeuGlnThrLeuGlnThr11eThrAsnLysLeuLys 954
 Db 2222 GAATGGAAATATCAGACAGACTCTTCAGACATTTGAAACTATCAAAATTAACCTAAA 2281
 Qy 955 ArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAlaSerGlu11eLeu11eAla 974
 Db 2282 AGGACTCTCAACAAAGAACCCCAATGATTCCTTTCAGCTTCATCAGAAATATCTATAGCC 2341

QY 975 AspSerAsnSerLeuGluThrLysLysAlaSerProPheCysArgProGlySerValLeu 994
 DB 2342 GACGACATTTCTATTAGAAACAAGAGCTTCCCTTCTGACAGCCAGCTCAGTGGC 2401
 QY 995 ArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThr 1014
 DB 2402 AGAGGGCGTAAATGTGTCATATGCTTGGAGACCTATTATATCTGGAAACATTTCAAC 2461
 QY 1015 CysGluSerCysArgTLeuGlySerTyrGlnAspGluGluGlyLeuGluCysValLeu 1034
 DB 2462 TGTGAACCTGCGGATCGAATCTTATCAAGATGAAGAGGCGCACTTGAAGTCAAGCTT 2521
 QY 1035 CysProSerGlyMetTyrThrGluTyrLLeuHisSerArgAsnLLeuSerAspCysValAla 1054
 DB 2522 TGCCCTCTGGGATGTATACGGAATATATCCATTCAAGAAACATCTCTGATTGTAAAGCT 2581
 QY 1055 GlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeu 1074
 DB 2582 CAGGTAAACAAAGGACCTACTACAGCAGTGAAGCTTGAACTTGTAATCTGTCCACTG 2641
 QY 1075 GlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
 DB 2642 GGCACCTTACGCGCAAAATTTGGTCCGAGCTGCTCTGTGTCAGAAACACCTCA 2701
 QY 1095 ThrValLysArgGlyAlaValAsnLLeuSerAlaCysGlyValProCysProGluGlyLys 1114
 DB 2702 ACTGGAAAGAGGAGCGCGTGAACATTTCTGATGATGAGGATTTCTGTCCAGAAAGAAA 2761
 QY 1115 PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsn 1134
 DB 2762 TTCTGCGCTTGGGGTAAATGCGCTGTCAACCAATGCTCTGTGATCAATTAACCAACTAAT 2821
 QY 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySer 1154
 DB 2822 GCAGGAGAGCGCTTCTGCGCTGCTGCTCTTTTATGAGAACCTACCAATCCGCTGCTTC 2881
 QY 1155 ArgSerLLeuThrGlnCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerVal 1174
 DB 2882 AGATCATCAAGAAATGTTCAAGTTTATGATCACTTTCTCAGCGGAGAGGAAAGCTGTG 2941
 QY 1175 ValProProAlaSerLeuGlyHisLLeuLysValArgHisGluLLeuSerGlnValPhe 1194
 DB 2942 GTGCCCCCTGCTCTTGTGACATATTAAGAGCAATTAAGCAATGCAAGTCAAGGTTTTC 3001
 QY 1195 HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGly 1214
 DB 3002 CATGAATGCTTCTTTAACTTGAACCAATAGTGAACCTGCGACGAACTTGGGCGCTGCT 3061
 QY 1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspLLeuAspGlu 1234
 DB 3062 TATGTTTGTCTGTCTCACTTGAATATACAGGCTTAAAGTGTGAACAGACATCGATGAG 3121
 QY 1235 CysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheLLeu 1254
 DB 3122 TGCAGCCCACTGCTTGTCAACATGAGATTTGTAAGACCTTATGCGGGAATTCAT 3181
 QY 1255 CysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnLLeuAsnGluCysSer 1274
 DB 3182 TGTGAATGCCCATCAGGTTTACACAGTCAACGCGTGTGAAGAAATTAATATGATGTAC 3241
 QY 1275 SerSerProCysLeuAsnLysGlyLLeuCysValAspGlyValAlaGlyTyrArgCysThr 1294
 DB 3242 TCCAGTCTTGTATTAAATTAAGGAATCTGTGTGATGTGTGTGCTGCTATCTTCAACA 3301
 QY 1295 CysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsn 1314
 DB 3302 TGTGGAAGAGATTTGTAGCGCTGCAATTTGAAGAGAAAGTCAATTAATGCCCATCAAC 3361
 QY 1315 ProCysLeuAsnAsnAlaValCysGluAspGlnValGlyPheLeuCysLysCysPro 1334
 DB 3362 CCATGCTTAAATTAATGCACTCTGTGAAGACAGAGTTGGGGATTTCTTGTGCAAAATGCCA 3421
 QY 1335 ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCys 1354

DB 3422 CCTGATTTTGGGTACCGAGTGGAAAGAAAGCTGACAGAGTGTCACTCAGCCATGCG 3481
 QY 1355 LysAsnGlyValAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysValAlaGly 1374
 DB 3482 AAAAATGAGCTTACCTGTAAACAGGTGCAATATGATTTCAAGTCCGTGTGTGAGCTGGC 3541
 QY 1375 PheThrGlySerHisCysGluLeuAsnLLeuAsnGluCysGlnSerAsnProCysArgAsn 1394
 DB 3542 TTCAAGAGTACACCTGTGAATTTGAACATCATGATATGTCAGTCTAATCACTAGAAAT 3601
 QY 1395 GlnAlaThrCysValAspGlyLeuAsnSerTyrSerCysLysCysGlnProGlyPheSer 1414
 DB 3602 CAGGCACTGTGTGATTAATTAATTCATTCAGTTGTAATGTCAAGCAGCAATTTCA 3661
 QY 1415 GlyLysArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
 DB 3662 GGCMAAGGTGTGAACAGAACAGTCTACAGGCTTTAACCTGGAATTTGAAGTTTCTGGC 3721
 QY 1435 LLeuTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
 DB 3722 ATCTATGATATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3781
 QY 1455 PheTyrMetLysSerSerAspAspMetAsnTyrGlyThrProLLeuSerTyrAlaValAsp 1474
 DB 3782 TTCTGATGAATATCCTCTGACGACATGAACATGAAGAACCAATCTCTATGACATGAT 3841
 QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLLeuTyrValAsn 1494
 DB 3842 AACGGACGACAAATCTTGTCTGACATGATTAACAGGCTGCTTCTTTATGTCAAT 3901
 QY 1495 GlyArgGlyLysLLeuThrAsnCysProSerValAsnAspGlyArgTyrHisLLeuAla 1514
 DB 3902 GCGAGGAAAGATTAACAATCTCCCTCGTGAATGATGACAGATGACATATATGCA 3961
 QY 1515 LLeuThrTyrThrSerAlaAsnGlyLLeuTyrLysValTyrLLeuAspGlyLysLeuSerAsp 1534
 DB 3962 ATCACTTGACAAATGCTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 4021
 QY 1535 GlyValAlaGlyLeuSerValGlyLeuProLLeuProGly 1547
 DB 4022 GGTGTGCTGTGCTCTCTGT 4080

RESULT 9
 BD158659 5124 bp DNA linear PAT 17-JAN-2003
 LOCUS
 DEPOSITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 BUKAYOKA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 5124)
 Ota,T., Iwaga,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,T.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 13502 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 FN JP 2002191363-A/13502
 PD 09-JUL-2002
 PI 28-JUL-2000 JP 2000280990
 PT TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
 PI SAITO
 PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
 PI KENICHI NAGAI, TETSUO OTSUKI
 PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/PC
 10, C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
 Primer for synthesizing full-length cDNA and use thereof FH Key

FEATURES
source
FT CMS Location/Qualifiers
(119) . (4066)
1. . 5124
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ALIGNMENT SCORES:

Alignment Scores:
Pred. No.: 0 Length: 5124
Score: 7537.00 Matches: 1347
Percent Similarity: 99.63% Conservative: 1
Best Local Similarity: 99.56% Mismatches: 5
Query Match: 73.75% Indels: 0
DB: Gaps: 0

US-09-977-053-6 (1-1842) x BD158659 (1-5124)

195 ThrAspGlyTyrSerAsnGlyIAspProArgProIleAlaAspSerLeuArgAspSer 214
2 ACTGATGATATTCATTCATGAGGAGACCTAGACCAATTCAGAGTCACGCGAGATTCA 61
215 GlyValGluIlePheThrPheGlyIleTrrGlnGlyAsnIleArgGluLeuAspMet 234
62 GAGGTGAGATCTTCATCTTGGCATATGGCAAGGAACTTCAGAGCTGAATGACATG 121
235 AlaSerThrProIleGluGluIleGlySerLeuLeuIleSerPheGluGluPheGluAla 254
122 GCTTCACCCCAAGAGAGAGACCTGTACCTGTACACAGTTTGAAAGATTGAGGCT 181
255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGluAspAspMet 274
182 TTAGCTGCGCGGCAATTCAGTAAGATCTACCTTGGAGTTTATTCAAGATGATATG 241
275 ValHisCysSerTyrLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu 294
242 GTCCACTGCTCATATCTTGTGTGATGAAAGGCAAGACCTGCTGTACCGAAATGGAGCTGC 301
295 LysCysGlyTyrHisThrGlyHisPheGluCysIleCysGluGluGlyTyrTyrGlyLys 314
302 AATATGGGACACACACAGGCCATTGTTGATGCACTGTGAAAGGGGATTACGGGAAA 361
315 GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro 334
362 GGTCTCAGATATGATGACAGCTTGCCTGCAATCGGAGCATACAAACCTGAGGCTCACCA 421
335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGlySer 354
422 GAGGAAATCAGCAGGTGCATTCATGTCGATGAAATATCACCTCTCCACCTGGAAAGC 481
355 ThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGlu 374
482 ACATCCCTCGAAGACCTGTCTGTCAAGAGGATACAGGCACTGTGCCAGACCTGTGAA 541
375 LeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGluAsnThrCys 394
542 CTGTGTCACTGCGCTGCGCTGGAAGCTCCGAAATAAGTTACTTATCCAAAACACTTGC 601
395 AsnAsnHisPhePheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
602 AACAAACCACTTCATGACCTGTGGGTCCGAGATCACCTGTGATTGATCTTGTGGAA 661
415 SerSerIleIleLeuGluCysLeuProAsnGlyLeuTrrSerGlySerGluSerTyrCysArg 434
662 AGCAGCATCATCTTATGTCTACCAACATGTTTGTGTGCGGTTTAAAGAGCTACTGAGA 721
435 ValArgThrCysProHisLeuAlaGlnProLysHisGlyHisIleSerCysSerThrArg 454
722 GTTAAAGAACTGTCTGATCTCCGCAAGCCGAAACATGGCCACATCACTGTCTTACAAAG 781
455 GluMetLeuTyrTyrSerThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly 474

782 GAATGTATATATAGACACATGTTGTTGCTGTGATGAGGATACAGTAAAGGC 841
475 SerAspLysLeuThrCysGlnGlyAsnSerGlnTrrAspGlyProGluProArgCysVal 494
842 AGTATAGACTTACTTGTCAAGAGAAACAGCCAGTGGATGAGGCAAGAACCCGGCTGTGTG 901
495 GluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCys 514
902 GAGCGCACATGTTCCACTTTCAGATGCCCAAGATTCATATATCCCAACATCTGT 961
515 GlyLeuGlnProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIle 534
962 GCGAAGACCAAGCAAAATTTGGACATCTGTATTAAGTTCCCGCAAGGTTCAATT 1021
535 LeuSerGlyValIlysgluMetLeuArgCysThrThrSerGlyLysTrrPheAsnValGlyVal 554
1022 TTATCTGAGTCAAGAAATGCTGATGATGATACCACTTCTGGAATAATGAAATGTGAGTT 1081
555 GluAlaAlaValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGlu 574
1082 CAGCAGACTGTGTATTAAGAGTGAAGCTTCCTCAATCAACTGTCTTAAAGACATTAAG 1141
575 AlaTyrThrLeuGlnGlnGlnAspSerAlaAsnValThrTrrGlnIleProThrAlaLys 594
1142 GCTTAAGCTTGGAAACAGAGATTTCTGCCATGTATCTGGCAGATTCCAAAGCTTAA 1201
595 AspAsnSerGlyGluGlyValSerValHisValHisProAlaPheThrProTrrLeu 614
1202 GACAACTGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1261
615 PheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGluAla 634
1262 TTCCCAATGAGATGTTGCTGATCTGATACCGCACTGACCTGATCCGGAACCAAGGCC 1321
635 SerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTrrCysArg 654
1322 AGCTGCATTTTCCATATCAAGGTTTATGATCCAAACCACTGTCTATGATGTGTGAGA 1381
655 SerProProProValGlnValSerGluLysValHisAlaAlaSerTrrAspGluProGln 674
1382 TCTCCACTCCCGCCAGGCTCCGAGAAAGATACATGCCGAACTGGAGATGAGCTCAG 1441
675 PheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu 694
1442 TTCTCAACACAACTCAGGAGCTGAATGTGTATACCAAGATCTCAACAGAGAACCTT 1501
695 PheProGlnGlyGluTrrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArg 714
1502 TTCCCTCAAGGAGAGACTATATGATACATATACGGCCACTGACCCCTCAGGCAATACAG 1561
715 ThrCysAspIleHisIleValIleLysGlySerProCysGluIleProPheThrProVal 734
1562 ACATGTATATCCATATTTGATCAATAAAGTTCTCCGTGAAATTCCATTCACACTGTGA 1621
735 AsnGlyAspPheIleCysThrProAspAsnThrCysValAsnCysThrLeuThrCysLeu 754
1622 AATGGGATTTTATATATCACTCCAGATATATCTGAGTCAACTGATCACTTACTGCTTG 1681
755 GluGlyTyrArgPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGly 774
1682 GAGGCGTATGATTTCAAGAAAGGCTCTACTGACAAAGTATTTATGTGCTTATGAAGATGC 1741
775 ValTrrLysProThrTyrThrThrGluTrrProAspCysAlaLysLysPheAlaAsn 794
1742 GTCTGAAACCAACATATATACCACTGAATGCGCAACCTGTCCCAAAAACGTTTTCAAAC 1801
795 HisGlyLysPheSerPheGluMetPheTrrLysAlaAlaArgCysAspAspThrAspLeu 814
1802 CAGGAGTTCAAGTCTTGAATGATTTCTACAAACAGCTCTGTGATGACAGATATG 1861
815 MetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCys 834
1862 ATGAAAGAGTTTCTGAAGCATTTTGAGACAGACCTGGGAAAAAATGTTCCCATCTTGT 1921

QY 835 SerAspAlaGluAspIleAspCysArgLeuGluGluAsnLeuThrIleValSerTyrCysLeu 854
DB AGTGTGCAAGAGCACTTGACCTGAGACCTGAGAGACCTGACCAAAAATATTCCTTA 1981
QY 855 GluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAla 874
DB GAATTAATTAATGACTATGAAAATGGCTTTGCAATTGACCAAGTGGCTGGGGTGCAGCT 2041
QY 875 AsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSer 894
DB AATAGGCTGAGATTACTCTTACCATGACTTCCTGACACATGTCACAAAACAGCCACAC 2101
QY 895 IleGlyAsnAlaIleSerSerArgIleIleValArgSerAlaProLeuSerAspTyrIle 914
DB ATCGGCAATCCCAAGCTCTCAGGATTAAAGAGAGCCCAATTAATGACTTAATTAAT 2161
QY 915 LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeu 934
DB AAGTTAATTTTAACATCAACAGCTAGTGGCCATTACCCGATGAAAGAAATGATACCTT 2221
QY 935 GluTyrGluAsnGlnGluArgLeuLeuGluThrIleThrIleAsnIleLeuVal 954
DB GAATGGGAAAATCAGACAGACTCTCTCAGACATGGAATCAATCAAAATAACTGAAA 2281
QY 955 ArgThrLeuAsnIleAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAla 974
DB AGACCTCTCAAGAACCCCAATGATTCCTTCACTTGCATCAGAAATCTTAATGCC 2341
QY 975 AspSerAsnSerLeuGluThrIleLysIleValAspProPheCysArgProGlySerValLeu 994
DB GACACCAATTCATTAAGAAACAAAAGGCTTCCCTTCGACGACCAAGGCTCAGTGTG 2401
QY 995 ArgGlyArgMetCysValIleAsnCysProLeuGlyThrTyrTyrAsnLeuGluIlePheThr 1014
DB AGAGGGCGATGTGTGTCATATGCCCCCTTGGAACTTAATTAATCTGAAACATTTCA 2402
QY 1015 CysGluSerCysArgIleGlySerTyrGluAspGluGluGlyGluLeuGluCysIleLeu 1034
DB TGTGAAAGCTCCGAGTCCGATCCCTTCAAGATGAAAGAGGCACTTAAGTGCAGCTT 2521
QY 1035 CysProSerGlyMetTyrThrIleGlyTyrIleIleIleSerArgAsnIleSerAspCysIleVal 1054
DB TCCCTCTCGGATGATACAGGAATATATCCATTCAGAAACATCTCTAATGTAAGCT 2581
QY 1055 GlnCysIleGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeu 1074
DB CAGTGTAAACAGGACCTACTCATGCAAGTGAAGCTTGTGAATGTGTCACTG 2641
QY 1075 GlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
DB GGACCTTATCAACCAAAATTTGGTTCCCGAGCTGCTCTCGTGTCAAGAAAACACTCA 2701
QY 1095 ThrValIleValArgGlyAlaValAlaAsnIleSerAlaCysGlyValProCysProGluGlyVal 1114
DB ACTGTGAAAAGAGAGCCGTGAACATTTTCATGATGAGGTTCCTGTCCAGAGAGAAA 2761
QY 1115 PheSerArgSerGlyLeuMetProCysIleProCysProArgAspTyrTyrGlnProAsn 1134
DB TTCTGGCGTTCTGGGTAAATGCTGTCAACCAATGCTCGTGAATTAATCAACCTAAT 2821
QY 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySer 1154
DB GCGAGGAAAGGCTTCTGCTGCTGCTCTCTTTTATGAACTACCCCAATTCGTGTTCC 2881
QY 1155 ArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerVal 1174
DB AGATCCATCAAGAAATGTTCAAGTTTACTTCAACTTCTCAGCGCAGAGGAAATGTGT 2941
QY 1175 ValProProAlaSerLeuGlyIleIleLysIleValArgIleGluIleSerSerGlnValPhe 1194
DB GTGCCCCCTGCTCTTGTGAATATTAAGAGGCAATGAATCAAGAGTCAAGTGTTC 2942

QY 1195 HisGluCysPhePheAsnProCysIleAsnSerGlyThrCysGlnGluLeuGlyArgGly 1214
DB CATGAATGCTTCTTAACTTGGCCCAATATGGAACCTGACGCAACTTGGCGGTG 3061
QY 1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLeuCysGluThrAspIleAspGlu 1234
DB TATGTTGTCTGTGTCACTTGATTAACAGGCTTAAGGTGAACAGCATCGAGTAC 3121
QY 1235 CysSerProLeuProCysLeuAsnAlaGlyValCysIleAspLeuValGlyIlePheIle 1254
DB TGCAACCACTGCTTCTGCTCTCAACATGAGATTGTAAAGCTTAATGGGAAATCAT 3181
QY 1255 CysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSer 1274
DB TGTGAGTGCATCAGGTTACACAGCTCAGCGGTGAAAGAAATTAATTAATGATGAC 3241
QY 1275 SerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThr 1294
DB TCCAGTCTTGTTTAATTAAGAAATGTGTGATGAGTGTGCTGCTATTCGTGCACA 3301
QY 1295 CysValIleGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsn 1314
DB TGTGTAAAGAAATTTGTAGGCTGCAATGTGAACAGAACTCAATATGCCAGTCAAC 3361
QY 1315 ProCysLeuAsnAlaValCysGluAspGlnValGlyIlePheLeuCysIleCysPro 1334
DB CCATGCTTAATTAATCAAGCTGTGAAGACAGGTTGGGGGATTCCTGTGCAATGCCCA 3421
QY 1335 ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCys 1354
DB CCGAGTTTTTGTGGTACCCGATGTGGAAGAAAGTGTGAATGATGTCTAGTCAAGCAATG 3481
QY 1355 LysAsnGlyAlaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGly 1374
DB AAAATGAGCTACTGTTAAGAGGCTGCCATAGCTTCAAGTCCGTGTGACAGTGGC 3541
QY 1375 PheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsn 1394
DB TTCAACAGATCACTGTGAATTAAGAACTCAATGAATGATCAAGTCAATCAATGAAT 3601
QY 1395 GlnAlaThrCysValAspGluLeuAsnSerTyrSerCysIleGlyGlnProGlyPheSer 1414
DB CAGGCCACCTGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3661
QY 1415 GlyGlnArgCysGluThrGlnGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
DB GGCAAAAGGTGTGAACAGAACTTACAGCTTAACCTGATTTGAAAGTTTCGCG 3721
QY 1435 IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
DB ATCTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3781
QY 1455 PheThrMetLysSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAsp 1474
DB TTCTGATTAATTCCTTAACAGATGATGATGATGATGATGATGATGATGATGATGAT 3841
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
DB AACGAGAGCAATACCTTGTCTCCAGTATTAAGAGGCTGGGTCTTATATGAT 3901
QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyLysArgTyrHisIleAla 1514
DB GCGAGGAAAAGATTAACAACTGTCCCTCGTAAAGATGCGAGAGGCAATATATGCA 3961
QY 1515 IleThrThrThrSerAlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAsp 1534
DB ATCACTTGACAGATGCCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 4021
QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
DB GGTGTGCTGGCTCTCTGTGTGTTGCCAATACCTGCT 4060

RESUL10

AK027870	LOCUS	AK027870	5124 bp	mRNA	linear	PII 01-AUG-2002
DEFINITION		Homo sapiens CDNA FL114964 fis, clone PLACE400581, moderately similar to FIBROPELIN 1 PRECURSOR.				
ACCESSION		AK027870				
VERSION		AK027870.1	GI:14042858			
KEYWORDS		oligo capping; fis (full insert sequence).				
SOURCE		Homo sapiens (human)				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS		1 Isogai,T., Ota,T., Hayaehi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Aotsuka,S., Yoshikawa,Y., Matsumawa,H., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Sasaki,N.				
TITLE		NEDD human cDNA sequencing project				
JOURNAL		Unpublished				
REFERENCE		2 (bases 1 to 5124)				
AUTHORS		Isogai,T. and Otsuki,T.				
TITLE		Direct Submission				
JOURNAL		Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)				
COMMENT		NEDD human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.				
FEATURES		Source				
source		1..5124				
		/location/Qualifiers				
		1..5124				
		/organism="Homo sapiens"				
		/mol_type="mRNA"				
		/db_xref="taxon:9606"				
		/clone="PLACE400581"				
		/tissue_type="placenta"				
		/clone_1ib="PLACE4"				
		/note="Cloning vector: pME186PL3"				
		119..4069				
		/note="unnamed protein product"				
		/codon_start=1				
		/protein_id="BAB55420.1"				
		/db_xref="GI:14042859"				
		/translation="MASTPKBKHCHYLHSPFEPBALARRALHEDLPFGSGPIODMWHC SYLDGKQCCDRMGSCKGTHTGECICERGVYKGLQYECTACPSGTYRBSGPG GHSGLPCPDENHTSPGSPEDPCVCRGVASGQTCILVHCPALKPENGFIONT CMHNNAACVYRCHGEPDVGSSITLCPNGMAGLESTCRPTGCHLRQPRHGLST STREMLYKTCVACDSEGRLEGSBDKLTQGSQMDGPRPCVBERHSTFPMKQVIT SPNKGKOPAKGTTI CVYSCQGFILSGVEMLRCTTSKMNMGVQAAVCAQVBAQI NCPKDIKATLEQODSANTWQIPTAKDMSGEKSVHAPLPTPLPIGVAIAYTT ATDLGNOASCFILIKVIDAEPPVDMGSGPPVQVSKVHNASMEPQFSDNAGL VTTBSHTGQDLPQGETTIVQATPDSPGNNRCDIHLVTKSGPCRIPTPVNGDICT PNTGVNCTLTLCBGDPTBEGSTDKTYKCAIBGVKMPYTTTTPPCAKKRPANHGKS FEMFKYAKACDDTDLAKKFSKAPFTLLGAVSPSCSDADIDMRLENLTKYCLATN YDVENGFALPGMGANLADYSYDFLDTQVBTATISGNASSRISKRAPIASDYIK LIFNTASVPLPDERNDTLEMNQORLLQTLTITNKLRTTNKLPMPYFQIASBLLI ADSNLSITKKAAPPKRGSVLGRMCCVPLGTYNVLHFTCESCRISGQDSQDLB CRLCSGWATTVYHISRNISIDKCAKCKGTVSGSGLETSGLTQAPKPPRSCISCS PENTSVRGANVYISACGVRCPBGRKSRSGMLPCHRCPRDYTPRANAKAPCLACPYG TTPFGRSRTITRCSSTFSAAERSVVPASIVGHTKGRHISQVFEHCFRPNCHNS GTCQDLGNGVYCLCPAGYTGKLCERTIDBCSPCLANNVCADLVGBRPICBPQGTG ORCBENINBSSPCILNKGI CVYGVAGYACTCVKGVGLACETRVNRCOSNPLNNAV CEDVGGFLCKCPGPIAGTCGNVDECLSGPCKNATCKDGNAPRCLCAAGFTSH CEININBOSNPNCRMDATCVDELINYSKCGQPSGKRCRTSGSTFENLDFVSGIYG YTMVGMKPSLHALCTPMKSSDDMWVCTPISYANDNSDNTLLITDNGVLYANG RKLTNCPSVNDGRMHAIITWTSANGIKVYIIDGLSDGAGLSGLPLPKMF"				
ORIGIN						
Alignment Scores:		0	Length:	5124		
Pred. No.:		7537.00	Matches:	1347		
Score:						

	Percent Similarity:	99.63%	Conservative:	1
	Best Local Similarity:	99.56%	Mismatches:	5
	Query Match:	73.75%	Indels:	0
	DB:	9	Gaps:	0
	US-09-977-053-6 (1-1842) x AK027870 (1-5124)			
Qy	195	ThrAspGlyTyrSerAsnGlyIysAspProAlaIleAlaSerLeuAlaAspSer	214	
Db	2	ACTGATGATATATTCCTCAATGAGGAGACCTTAGACCAATTCAGGTACGCGAATTCA	61	
Qy	215	GIYValGluIlePheThrPheGlyIleTyrGluIYAsnIleArgGluLeuAspAspMet	234	
Db	62	GGAGTGGAGATCTTCACTTTGGCATATGCGAAGGACATTCGAGCTGAATGACATG	121	
Qy	235	AlaSerThrProIysGluGluIleIscYerTleuLeuIleSerPheGluGluIleGluAla	254	
Db	122	GCTCCACCCCAAGAGGAGGACCTGTATCCTGCTACACAGTTTGAAGAATTGAGGCT	181	
Qy	255	LeuAlaArgAlaLeuIleGluAspLeuProSerGlySerPheIleGluAspAspMet	274	
Db	182	TTAGCTCGCGGACATTCGATGAAAGATCTTCTGGAGTTTATTCAGATGATATG	241	
Qy	275	ValIleCysSerTyrLeuCysAspGluGlyIysAspCysCysAspAspMetGlySerCys	294	
Db	242	GTCATCTGCTCATATCTTGTGATGAAAGGACAGCTGTGACCGAATGGAGGCTGC	301	
Qy	295	IysCysGlyThrIleThrIleGlyIlePheGluCysIleCysGluIysGlyTyrTyrGlyIys	314	
Db	302	AAATGGGAGCACACACAGGCAATTTTGAAGTGTGATGATGATGATGATGATGATG	361	
Qy	315	GIYLeuGluIleTyrGluCysThrAlaCysProSerGlyThrTyrIysProGluIysSerPro	334	
Db	362	GGTCTCAGATATGAAAGGACAGCTTCCATCGGAGCAATACAAACCTGAAAGCTCACA	421	
Qy	335	GIYGIYIleSerSerCysIleProCysProAspGluIysAsnIleThrSerProGluIysSer	354	
Db	422	GGAGATCAGAGATTCATCTTCCATGCTGATGAAATACACCTTCCACCTGAGAAC	481	
Qy	355	ThrSerProGluIysAspCysValCysArgGluIYIYArgAlaSerGlyGluIleThrCysGlu	374	
Db	482	ACATCCCTGAAAGACGTGTCTGACAGAGGAGATACAGGCACTGGCAGACCTGAGAA	541	
Qy	375	LeuValIleCysProAlaLeuIysProProGluIysAsnGlyTyrPheIleGluAsnThrCys	394	
Db	542	CTGTGCACCTGACCTTCCCTGAGAGCTCCGAAATGGTACTTATCCAAACACACTGC	601	
Qy	395	AsnAsnIlePheAsnAlaIleCysGlyValArgCysHisProGlyPheAspLeuValGly	414	
Db	602	AAACACACCTTCAATGACGCTGTGGGATCCATGACCTGATTTGATTTGTGGGA	661	
Qy	415	SerSerIleIleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArg	434	
Db	662	AGAGCATATCTTATGTGTACCAATGTTGTGGTGGGTTTGAAGAGCTACTGACAGA	721	
Qy	435	ValArgThrCysProIleIleLeuArgGluProIysHisGlyIleIleSerCysSerThrArg	454	
Db	722	GTAAGAACATGCTCTCATCTCCCGCAGCGGAAACATGCGCACATCAGCTGTCTACAGG	781	
Qy	455	GluMetLeuTyrTyrThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly	474	
Db	782	GAATATGATATTAAGAACATATTTGGTTGCTGTGATGAAAGGATACATGAGAGGC	841	
Qy	475	SerAspIysLeuThrCysGluGlyIysAsnSerGluTyrAspGlyProGluProArgCysVal	494	
Db	842	AGATATGAGCTTACTTGTCAAGGAAACAGCAGTGGAGATGGCCAGAACCCCGGTGTGTG	901	
Qy	495	GluArgHisCysSerThrPheGluIleCysProIysAspValIleIleSerProHisAsnCys	514	
Db	902	GAGCGCCACTGTTCCACTTTCAGATGCCCAAGATGTCATCATCCCCCAACACATCTG	961	
Qy	515	GIYLeuGluProAlaIysPheGlyThrIleCysTyrValSerCysArgGluIlePheIle	534	

Db 962 GGCAGACGACGACCAAAATTTGGAGCGATCTGCTATGTAAGTTCGCGCCAGGCTTCATT 1021
Qy LeuSerGlyValIleGlyMetLeuArgCysThrThrSerGlyValSerThrPheValGlyVal 554
Db 1022 TTATCTGGAGTCAGAAAGAAAGCTGAGATGTACCACTTCGGAAAGATGGATTCGGAGTT 1081
Qy 555 GlnAlaAlaValCysLeuAspValGluAlaProGlnIleAsnCysProLysAspIleGlu 574
Db 1082 CAGGAGCTGTGTGTGTAAGAGCTGGAGGCTCTCAATCACTGTCTTAAGGACATAGAG 1141
Qy 575 AlaValThrLeuGlnGlnIleAspSerAlaAsnValThrTrpGlnIleProThrAlaLys 594
Db 1142 GCTAAGACTCTGGAAACAGCAAGATTCGCAATGTACTCGCAAGATTCCAACAGCTAAA 1201
Qy 595 AspAsnSerGlyGluValSerValHisValHisProAlaPheThrProProGlyLeu 614
Db 1202 GACCACTCTGGGAAAGGTGTCAGTCCAGTTCATCCAGCTTCACCCCACTTAACCTT 1261
Qy 615 PheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyValAsnAla 634
Db 1262 TTCCCAATTTGGAGATGTTGCTATCGTATACAGGCACTGACTATCCGGCAACGAGGCC 1321
Qy 635 SerCysIleIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArg 654
Db 1322 AGCTGCATTTTCCATATCAAGGCTTATGTATGACAGAACCACTGTCTATACACTGTGCAGA 1381
Qy 655 SerProProProValGlnValSerGlyLysValHisAlaAlaSerTrpAspGluProGln 674
Db 1382 TCTCCACCTCCCGTCAGAGGTCTCGAGAAAGGTACATCCGCAAGCTGGAGATGAGCTCAG 1441
Qy 675 PheSerAspAsnSerGlyValaGluLeuValIleThrArgSerHisThrGlnIleAspLeu 694
Db 1442 TTCTCAGCACTCAGGGGCTAAATTGGCTATTAACAGAAATCATACACAGAGAGCTTT 1501
Qy 695 PheProGlnGlyGlnThrIleValGlnTyrThrAlaThrAspProSerGlyValAsnAlaArg 714
Db 1502 TTCCCTCAGGGGAGACTATAGTACAGTATACGCGCACTGACCCCTCAGGCAATACAGAG 1561
Qy 715 ThrCysAspIleHisIleValIleLysGlySerProCysGluIleProPheThrProVal 734
Db 1562 ACATGTATATCCATATGTCTATAAAGGTTCCTCCGTGAAATTCATTCACCTGTGA 1621
Qy 735 AsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeu 754
Db 1622 AATGGGATTTTATATGACATCCAGATATACTGGAGTCACTGATTAACCTTCTTG 1681
Qy 755 GlnGlyTyrAspPheThrGlnGlySerThrAspLysTyrTyrCysAlaTyrGluAspGly 774
Db 1682 GAGGCTATGTATTCACAGAAAGGCTCTACTGACAGATATTAATGTATATGAAAGTGGC 1741
Qy 775 ValTyrLysProThrTyrThrThrGlnTyrProAspCysAlaLysLysArgPheAlaAsn 794
Db 1742 GTCTGGAAACCAACATATACCTGATATGGCCAGACTGTGCCAAAACGTTTTCGAAC 1801
Qy 795 HisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeu 814
Db 1802 CACGGGTTCAAGTCTTTGAGATGTTCTACAAAGCACTCGTTTGATGACACAGATCTG 1861
Qy 815 MetLysLysPheSerGluAlaPheGlnThrThrLeuGlyLysMetValProSerPheCys 834
Db 1862 ATGAAGAAAGTTTTCGAAGCATTTGAGAGACCTTGAGAAAATGTCTCCATCATTTGT 1921
Qy 835 SerAspAlaGluAspIleAspCysArgLeuGlnGluAsnLeuThrLysLysTyrCysLeu 854
Db 1922 AGTATATGAGAGACATTTGACCTGAGACTGAGAGACCTGACCAAAAATATATGGCTA 1981
Qy 855 GlnTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAla 874
Db 1982 GAATATATATATGACTATGAAATATGCTTTGCAATTTGACAGAGTGGCTGGGTGAGCT 2041
Qy 875 AsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGlnGlnThrAlaThrSer 894
Db 2042 AATAGGCTGAGTACTTATCGATGACTCTGAGCACTGTGCAAGAAACAGCCACACAGCC 2101

Qy 895 IleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIle 914
Db 2102 ATGGGCATNGCCAAAGTCTCCAGGATTAAGAAGAGGCCCACTTATCTACATATAAATT 2161
Qy 915 LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluLysAsnAspThrLeu 934
Db 2162 AAGTTAATTTTAAATCAACACCTAGTGTGCATTAACCGAATGAAAGAAATATACCTT 2221
Qy 935 GlnTrpGluAsnGlnIleArgLeuLeuGlnThrLeuGlnIleThrAsnLysLeuLys 954
Db 2222 GAATGGAAAATACAGACAGACTCTTCAGACATTTGGAACTATACAAATTAACCTGAA 2281
Qy 955 ArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAla 974
Db 2282 AGGACTCTCAACAAAGACCCCATGATTCCTTACGCTTGCAATCAAAATATCTTAAGCC 2341
Qy 975 AspSerAsnSerLeuGlnThrLysLysAlaSerProPheCysArgProGlySerValIleu 994
Db 2342 GACAGCAATTCATTAGGAACAAAAAGGCTTCCCTTCGACAGACAGGCTCAGTGTG 2401
Qy 995 ArgGlyArgMetCysValAsnCysProLeuGlyTyrTyrTyrAsnLeuGlnHisPheThr 1014
Db 2402 AAGAGGCGTATGTGTCAATTCCTTTGGAACTATTAATGTGAACATTTCAACC 2461
Qy 1015 CysGluSerCysArgIleGlySerTyrGlnAspGlnGlyGlnLeuGlnCysLysLeu 1034
Db 2462 TGTGAAGCTGCCTGAGTGGATCTCATCAAGATGAAGAGGCAACTTGATGTCAGACTT 2521
Qy 1035 CysProSerGlyMetTyrThrGlnTyrIleHisSerArgAsnIleSerAspCysLysVala 1054
Db 2522 TCCCTCCCTGGAGATGACAGAAATATATCCATTCAGAAACATCTCAATGTATAAGCT 2581
Qy 1055 GlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGlnThrCysGluSerCysProLeu 1074
Db 2582 CAGTGTAAACAGGACCTACTCAAGCACTGAGACTTGACATTTGTAATCTGTCACTG 2641
Qy 1075 GlnTyrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
Db 2642 GGCATTAATCAGCCAAATATTTGGTCCCGGACTGCTCTCGGTCCAGAAAACACTCA 2701
Qy 1095 ThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGlnGlyLys 1114
Db 2702 ACTGTGAAGAAAGAGGCCGTGAACATTTCTGATGTGAAGTTCTTGTCCAGAGGAAAB 2761
Qy 1115 PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsn 1134
Db 2762 TTCTGCGCTTCGGTTATATGCCCTGTCAACCATGTCTCTGTGACATTAACCACTTAAT 2821
Qy 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyTyrThrProPheAlaGlySer 1154
Db 2822 GCAGGGAAGGCTTCTGCTGCTGTGCTGTCTTTATGAACTAACCCCATTCGCTGTGCTC 2881
Qy 1155 ArgSerIleThrGlnCysSerSerPheSerSerThrPheSerAlaAlaGlnGluSerVal 1174
Db 2882 ACATCAATCAACGAATGTTCAGATTTAGTTAACTTTCACGGCAGAGGAAAGTGTG 2941
Qy 1175 ValProProAlaSerLeuGlyHisIleLysLysArgHisGlnIleSerSerGlnValPhe 1194
Db 2942 GTGCCCCCTGCTCTTGTGACATATTAAGAGGCAATGAATACAGACTGACGTTTTC 3001
Qy 1195 HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyValGly 1214
Db 3002 CATGAATGCTTCTTTAACTCTTGCCACATATGTGAACCTGCGACAGMACCTGGGCTGTG 3061
Qy 1215 TTTValCysLeuCysProLeuGlnTyrThrGlnLysLeuLysCysGlnThrAspIleAspGlu 1234
Db 3062 TATGTTTGTCTGTCTGCACTTGATATACAGGCTTAAGATGTGAACACACATCATGAG 3121
Qy 1235 CysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlnLysPheIle 1254
Db 3122 TGCAGCCCACTGCTTGTGCTCAACATGAGATTTGTAAAGACTAGTTGGGGAATTCATT 3181

Oy	1255	CysgluCyPProserGlyYTrthGlyGlnAArgCysgluGluAenIleAenGluCySer	1274
Db	3182	TGTGAGTCCCATCAGGTTCACAGGTCCAGCGGTGTGAGAAATATTAAGAGGTGAGC	3241
Oy	1275	SerSerProCySleuAenIleGlyYIleCyValAspGlyValAlaGlyTYrArgCyThr	1294
Db	3242	TCAGTCTCTTGTTTAATTAAGAAATCTGTGTGTAAGTGTGGCTGGCTTAATCTTGCA	3301
Oy	1295	CysValIleGlyPheValGlyLeuHisCysgluTrhGluValAsnGluCySglInserAsn	1314
Db	3302	TGTGTGAAGAAGATTGTGAGGCTGTCATTTGTGAACAGAAAGTCATTAAGATGCAAGTCAAC	3361
Oy	1315	ProCySleuAenAenAlaValCySgluAspGlnValGlyGlyPheLeuCySlySPro	1334
Db	3362	CCATCTCTTAATTAAGTCAAGTCTGTGACACCAAGTGTGGGGAATCTTGTGCAAAATGCCCA	3421
Oy	1335	ProGlyPheLeuGlyTYrArgCysGlyYlysAenValAspGluCySleuSerGluProCys	1354
Db	3422	CCTGATTTTTGGGATACCCGATGTGAAAGAACGTGATGATGTCTCAAGTCACGCAATGC	3481
Oy	1355	LYSAENGlyAlaThrCysIleAspGlyAlaAenSerPheArgCysLeuCyAlaAlaGly	1374
Db	3482	AAAAATGAGCTACTCTGTAAAGACGCTGCCAATAGCTTCAGATGCTGTGTGCACTGGC	3541
Oy	1375	PheTrhGlySerHisCysGluLeuAenIleAenGluCySglInserAsnProCyArgAsn	1394
Db	3542	TTCAACAGATCACACTGTGAATTTGAACATCAATGATGATGCTTAATCCATGTAGAAT	3601
Oy	1335	GlnAlaThrCyValAspGluLeuAenSerTYrSerCySlySProGluProGlyPheSer	1414
Db	3602	CAGGCACTGTGTGATGAATTAATTCATACAGTTGTAATGTACACCGAGATTTTCA	3661
Oy	1415	GlyGlnAArgCysGluTrhGluInserTrhGlyPheAenLeuAerPheGluValSerGly	1434
Db	3662	GGCANAAGGTGTGAACAGAACAGCTACAGGCTTTAACCCTGATTTGAAGTTCTTGCC	3721
Oy	1435	IleTYrGlyTYrValMetLeuAspGlyMetLeuProSerIleuHisAlaLeuThrCyThr	1454
Db	3722	ATCATGTGATATGTCACTGATGTTGGCAGTCCCATCTTCACAGCTCTAACCTGTACC	3781
Oy	1455	PheTrpMetIleSerSerAspAspMetAsnTYrGlyTYrThrProIleSerTYrAlaValAsp	1474
Db	3782	TTCTGGATGAATCTCTGACACATGAACTATGGAACACCAATCTCTATGCAAGTGTAT	3841
Oy	1475	AsnGlySerAspAenTrhLeuLeuLeuThrAspTYrAsnGlyTYrValIleuTYrValAsn	1494
Db	3842	AACGCAACGACAAATACCTTGCTCTGACTGATTTAAACGCTGGCTCTTAATGTAAAT	3901
Oy	1495	GlyArgGluYValIleTrhAsnCyBProSerValAenAspGlyArgTYrHisIleAla	1514
Db	3902	GGCAGGAGAAAGATTAACAACATGTCCTCGGAGATGATGGCAGATGCATCATATGCA	3961
Oy	1515	IleThrTYrThrSerAlaAsnGlyYIleTrpIleValTYrIleAspGlyYlysLeuSerAsp	1534
Db	3962	ATCATCTTGACAAAGTGCACATGCGCATCTGGAAGAGCTATATCGAAGGGAATATATGAC	4021
Oy	1535	GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly	1547
Db	4022	GGTGGTCTGGCTCTCTGTGGTTTGCCCATACCTGTAT	4060
RESULT 11			
AK122605 3253 bp mRNA linear PRI 09-SEP-2003			
LOCUS Homo sapiens cDNA FLJ16013 f.8, clone PLAC5000171, weakly similar			
DEFINITION to B-SELECTIN PRECURSOR.			
ACCESSION AK122605			
VERSION AK122605.1 GI:34527780			
KEYWORDS oligo cloning; f.8 (full insert sequence).			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE 1			

AUTHORS	Oshima, A., Takahashi-Pujila, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Mashimino, K., Yuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Matsumura, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahata, K., Masuko, Y., Nagai, K. and Iwaga, T.
TITLE	NEO human cDNA sequencing project
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 3253)
AUTHORS	Iwaga, T. and Yamamoto, J.
TITLE	Direct Submission
JOURNAL	Submitted (15-JUL-2003) Takao Iwaga, FLJ Project (HRI Team); 2-6-7 Kazuo-Kametani, Kisarazu, Chiba 292-0818, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES	Location/Qualifiers
SOURCE	1..3253 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="PLACES000171" /issue_type="placenta" /clone_id="PLACES" /note="Cloning vector: pME18SFL3"
ORIGIN	
Alignment Scores:	
Pred. No.:	1,12e-303 Length: 3253
Score:	4776.00 Matches: 866
Percent Similarity:	99.88% Conservative: 0
Best local Similarity:	99.88% Mismatches: 1
Query Match:	46.74% Indels: 0
DB:	9 Gaps: 0
US-09-977-053-6 (1-1842) x AK122605 (1-3253)	
Qy	1 MetTTPProaTgLeuAlaPheCyGvETfTgLYLeuAlaLeuValSergLYTPAlaTthr 20
Db	334 ATGGGCGCTCGCCCTGGGCTTTTGTGTGGTGGGGGCTGGGCGCTCGTTTGGGCTGGGCGACCC 393
Qy	21 PheGlnGlnMetSerProSerATgAaMPhSerPheArgLeuPheProGlnTThrAlaPro 40
Db	394 TTTTCAGAGAAATGTCCTCCGTGGCGCAATTTAGCTTCGCTTCCCGGAGACCGCGCCC 453
Qy	41 GLYAlaProGlySerLeuProAlaProProAlaProGlyVasPGluAlaAlaGlySerArg 60
Db	454 GGGGCCCCCGGAGATATCCCGCGCGCCCGCGCTCTGGCGACAGAGCGGCGGAGACAGA 513
Qy	61 ValGluArgLeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGlnLeuSergLnu 80
Db	514 GTGAGAGCGGCTGGGCGCGAGGGTTCGGGCGAGCGGTGGCGGCTGCGCGAGAGCTCAGAGAG 573
Qy	81 ArgLeuGlnLeuValPheLeuValaAspAspSerSerValGlyGlyValaAspPheArg 100
Db	574 CGCTGGAGGCTGTGCTTCTGTGGAGAAATTCGTCCAGCGATGGGCGAGAGTCACTTCGCG 633
Qy	101 SerGlnLeuMetPheValArgLYSLeuLeuSerAspPheProValValProThrAlaThr 120
Db	634 AGGAGGCTCAATGTTTCGTCGCGCAAGCTGCTGTCCGACTTCCCGTGGTGGCCAGCGGACAG 693
Qy	121 ArgValAlaIleValThrPheSerSerLYSAsnTYrValValProArgValaAspTYrTle 140
Db	694 CGCGTGCCTATCGTGACCTTCTGTGTCCAAAGCTACGATGTGGCGCGCGTGCATTTTCATC 753

OY		141	SerThrArgAlaGlnIleHisGlyCysValAlaLeuLeuLeuGlnGluIleProAlaIle	160
Dd		754	TCCACCGGCCGGCGGCAGACAAGAAGCGCGCTCCCTCCAGAAATCCGTGCATC	813
OY		161	SerTyrArgGlyGlyGlyTThrTyrThrGlyGlyAlaPheGlnGlnAlaAlaGlnIleLeu	180
Dd		814	TCTTAACCAAGGTGGCGGCACTTCACCAAGGGCGCTTCAGCAAGCGGCMAATTCTT	873
OY		181	LeuHISalAargGluAsnSerThrlsValValPheLeuIleThrAspGlyTyrSerAsn	200
Dd		874	CTTCATGCTGAAGAAAATCAACAABAAAGTTGATTTCTCATCACTGAATGATATTCAT	933
OY		201	GlyGlyAspProArgProIleAlaAserLeuArgAspSerGlyValGluIlePheThr	220
Dd		934	GGGGGAGACCCTTAGACCAATTGACGCGTCACTGGAGATTACAGAGTGGAATCTTCACT	993
OY		221	PheGlyIleTPrgInglYAsnIIlaArgGluLeuLeuAsnAPMeAlaserThrProlysGlu	240
Dd		994	TTGGCATATGGCAAGGAAACATTCGAAGCTGAATTAACATGGCTTCCACCCCAAGGAG	1053
OY		241	GluHisCysTyrLeuLeuHISerPheGlnGluPheGlnAlaLeuAlaArgAlaAlau	260
Dd		1054	GAGCACCTTACTCTGTCTACACAGTTTGAAABAATTTGAGCGTTTAACTCGCGGGCANTG	1113
OY		261	HISgluAspLeuProSerGlySerPheIleGlnAspAPMeValHisCysSerTyrLeu	280
Dd		1114	CATGAAGATCTAACCTTCGGGAGTTTTTATTCAGATGATATGATGTCACCTGCATATCTT	1173
OY		281	CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyTThrHisPthr	300
Dd		1174	TGTATGAAAGGCAGAGACCTGCTGTACCGAATGGCACTCAATAATGGGACACACACA	1233
OY		301	GlyHisPheGlnCysAlaCysGlnLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys	320
Dd		1234	GGCCATTTTGAGTGCATCTGTGAABAAAGGGATTATACGGGAAAGCTGTGCATTATGAATGC	1293
OY		321	ThrAlaCysProSerGlyTThrTyrLysProGlnGlySerProGlyGlyLeSerSerCys	340
Dd		1294	ACAGCTTCCCATCGGGAGCAATACAACTGAAGGCTCACAGGAGGAATCAGACAGTTGC	1353
OY		341	IleProCysProAspGluAsnHisPthrSerProProGlySerThrserProGluAspCys	360
Dd		1354	AATCCATCTCTGAAGAAATCACACCTCTCCACTGGAMACATATCCCCGAAAGACTGT	1413
OY		361	ValCysArgGlnGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla	380
Dd		1414	GTTCTGCAAGAGGGATACAGGCGATCTGGCAGACCTGTGAATCTTGTCACTGCCCTGCC	1473
OY		381	LeuLysProProGluAsnGlyTyrPheIleGlnIleThrCysAsnAsnHisPheAsnAla	400
Dd		1474	CTGAAGCTCCCGAABAAATGGTTACTTATTCMAAACCTTCACACAAACCACTTCATGCA	1533
OY		401	AlaCysGlyValAlaArgCysHisProGlyLysAspLeuValGlySerSerIleIleLeuCys	420
Dd		1534	GCGTGTGGGGTCCAAATGCACCTGTGAATTTATCTTGGAGAAAGCAGATCATCTTAAGT	1593
OY		421	LeuProAsnGlyLeuTyrSerGlySerglusertYrCysArgValArgThrCysProHis	440
Dd		1594	CTATCCCAATGGTTGTGTGGTCCGGTTCAAGAGACTCTGCAAGATGAAGAACTATCTCAT	1653
OY		441	LeuArgGlnProLysHISegLyHisIleSerCysSerThrArgGluMetLeuTyrLysThr	460
Dd		1654	CTCGCGCAGCGGAACATGGCCACATCGCTGTCTTCAAGGGAATGTATATTAAGACA	1713
OY		461	ThrCysLeuValAlaCysAspGlnGlyTyrArgLeuGlnGlySerAspLysLeuThrCys	480
Dd		1714	ACATGTATTGGTGCCTGTGATGAAGAGGTACAGACTGAAGAGCAAGTGAAGCTTAACTTGT	1773
OY		481	GlnGlyAsnSerGlnTTrpAspGlyProGluProArgCysValGluArgHisCysSerThr	500
Dd		1774	CAGGAAACAGCGACTGGATGGGCACAGAACCCCGGTGTGTGAGCGGCACCTGTTCCACC	1833
OY		501	PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys	520

Db	1834	TTTGAGATGCCCAAGATGTCATCATATATCCCCCAACACTGTGGACAGCCAGCCAAA	1893
QY	521	PheGlyThrIleCytryValSerCyargGlnGlyPheIleLeuSerGlyValIleGlu	540
Db	1894	TTTGGAGCAATCTCTCATGTATGAATGGCCGCAAGGGGTTCATTTTATCTGGAGTCMAAGAA	1953
QY	541	MetLeuArgCyThrThrTrpSerGlyValSerPheValGlyValGlnIleAlaValCyIys	560
Db	1954	ATGCTGAGAGTGAACACTCTTCGAAAAATGAAATGTCCAGATTCAGGCAAGCTGTGTGAAA	2013
QY	561	AspValGlnAlaProGlnIleAsnCyProLysAspIleGlnAlaIleThrLeuGlnGln	580
Db	2014	GACGTGGAGGCTCTCTCAATCAATCTGCTTAAGGACATAGGGCTAAGGCTTGGACAG	2073
QY	581	GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlnLys	600
Db	2074	CAGAAATCTGCAATGTACTCTGGCAATTCACACACTAAGCAACCTGTGGTGAAGAAG	2133
QY	601	ValSerValHisValHisProAlaPheThrProProGlyLeuPheProIleGlyAspVal	620
Db	2134	GTGTGACGTCCAGTTTCAATCCAGCTTCAACCCCACTTACCTTTCCCAATTGGAGAGTT	2193
QY	621	AlaIleValIyThrAlaThrAspLeuSerGlyValGlnAlaSerGlyIlePheHisIle	640
Db	2194	GCTATCGATATCAAGGCAACTGACCTATCCGGCAACAGAGCCAGCTCATTTTCCAAATTC	2253
QY	641	LysValIleAspAlaGlnProProValIleAspTrpCyArgSerProProValGln	660
Db	2254	AAGGTATATGATGCAAGCAACCTGTATAGACTGCGCAAGATCTCCACTCCCGTCAG	2313
QY	661	ValSerGlnLysValHisAlaAlaSerTrpAspIleProGlnPheSerAspAsnSerGly	680
Db	2314	GTCTCGAGGAAGATACATGCGCAAGCTGGAGTAGCCTCAGTTCTCAGACAACTCCAGG	2373
QY	681	AlaGluLeuValIleThrArgSerHisThrGlnIleAspLeuPheProGlnIleGlnThr	700
Db	2374	GCTAAATTTGGTCATATACCAAGATGCAACACAAAGAAACCTTTTCCCTCAAGGGGAACT	2433
QY	701	IleValGlnIyThrAlaThrAspProSerGlyValAsnAsnArgThrCyAspIleHisIle	720
Db	2434	ATATGATCAGTATACGCCACTGACCCCTCAGGCAATTAACGAGCAATGATATCCATATT	2493
QY	721	ValIleLysGlySerProGlyIleProPheThrProValAsnGlyAspPheIleCys	740
Db	2494	GTCATTAAGGTTCTCCCTGTGAATTCATTCACACTGTAAATGGGATTTATATGTC	2553
QY	741	ThrProAspAsnThrGlyValAsnCyThrLeuThrCysLeuGlnGlyIyTrpAspPheThr	760
Db	2554	ACTCCAGATTAATCTGAGTCACACTGTACATTAATCTTGCTTGGAGGGCTATGATTTCCA	2613
QY	761	GlnGlySerThrAspLysIyTrpCyAlaIleTyrluAspGlyValIyTrpLysProThrTy	780
Db	2614	GAAAGGCTTACTGACAAAGTATATTATGTGCTTATGAAGATGGCGCTGAAACCAACTAT	2673
QY	781	ThrThrGlnTrpProAspCyAlaLysLysLysArgPheAlaAsnHisGlyPheLysSerPhe	800
Db	2674	ACCACTGAATGGCAGACTGTGCCAAAAACGTTTCCAAACCAACGAGGTCACAGCTCTT	2733
QY	801	GlnLeuSerPheTyIysAlaAlaArgCyAspAspThrAspLeuMetLysLysPheSerGlu	820
Db	2734	GAGATGTTCTTACAAAGCAGCTCGTGTGATGACACAAATCTGATGAAGAGTTTTCGAA	2793
QY	821	AlaPheGlnThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle	840
Db	2794	GCAATTTGAGAGCACTCGGGAAATATGTCCCAATTTGTATGATGCAAGAGCAACTT	2853
QY	841	AspCyAsnArgLeuGlnIleAsnLeuThrLysLysIyTrpCysLeuGlnIyTrpAsnIyTrp	860
Db	2854	GACTGCAAGCTGGAGAGAACTTACCAAAAAATATGCTAGAAATTAATTAATGACATAT	2913
QY	861	GluAsnGlyPheAlaIleGly 867	

Db 2914 GAAATGGCTTTGCAATTGGT 2934

RESULT 12

AX880735 1969 bp DNA linear PAT 17-DEC-2003

LOCUS Sequence 15640 from Patent EP1074617.

DEFINITION AX880735

ACCESSION AX880735

VERSION AX880735.1 GI:40035471

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.

TITLE Primers for synthesizing full-length cDNA and their use

JOURNAL Patent: EP 1074617-A 15640 07-FEB-2001;

RESEARCH Association for Biotechnology (JP)

FEATURES

source Location/Qualifiers

1..1969

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

415..1656

/note="unnamed protein product"

/codon_start=1

/protein_id="CA891063.1"

/db_xref="GI:40035472"

/translation="MLYKTLCLVACDGEYLEGSDKLTTCGNSQWDPGPRCVERHGS
TQMPKDVIIIPNCQKQPAKFGTICVSCROGPIILSGVKMLRCTTSGKNVGVQAA
VCKDRAPOINCPKDIKALTEQDDSANVTWQIPTAKDNGSKVSRVHPAPTPYLF
PIGDVALVATDLSGNQASCIPIHIVIDAEPVIDWCRSPFPVQVSEKVKHAAWDEP
QFSDNGAELVITRSHQCDLPQGETIYQVATDPSGNRTCDIHLVIGKSPCEIIP
TPVNGDFICTPDNTGNTCTCLEGDFTEGSDTKYCYATEDGVKFKPTTYTTEPDCNK
KRFANHGFKSPFMYAARCDTDLAKKFSLEALTTILGKVPSPFCSDAEDIDCLREN
LTKYCLVNYDYENGPAIGN"

ORIGIN

Alignment Scores: 3,55e-192 Length: 1969

Pred. No.: 3073.00 Matches: 547

Score: 3073.00

Percent Similarity: 99.45% Conservativeness: 0

Best Local Similarity: 99.45% Mismatches: 3

Query Match: 30.07% Indels: 0

DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x AX880735 (1-1969)

Qy 318 TyrGluCysThrAlaCysProSerGlyThrTyrlsProGluGlySerProGlyGly 337

Db 1 TATGATGCAACACTTGGCCATCGGGACATACAACTGAGCCTCACCAGGAGGATC 60

Qy 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357

Db 61 AGCAGTTGCATTCCATGTCGCGATGAAATCACACCTCTCCACCTGGAAGCAGATCCCT 120

Qy 358 GluAspCysValCysArgGluGlyTyrlsArgAlaSerGlyGlnThrCysGluLeuValHis 377

Db 121 GAAGACTGTGTCTGCGAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAATCTGTCAC 180

Qy 378 CysProAlaLeuLysProProGluAsnGlyTyrlsPheIleGlnAsnThrCysAsnAsnHis 397

Db 181 TGCCTGCTGCTGAGCCTCCCGAAATGGTTACTTTATCCAAACACTTGCACACACAC 240

Qy 398 PheAsnAlaAlaCysGlyValArgCysHisPheProGlyPheAspLeuValGlySerSerIle 417

Db 241 TTCAATGCGACCTGTGGGGTCCGATGTCACCTTGGATTTGATCTTGTGGGAAGCAGCATC 300

Qy 418 IleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerGluSerTyrlsArgValArgThr 437

Db 301 ATCTTATGCTCTCCCAATGGTTTGTGGTCCGCTTCAGAGAGCTACTGCAGAGTAAGAACA 360

Qy 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457

Db 361 TGTCTCTCATCTCCGCGAGCCGAAACATGGCCACATCAGCTGTCTCAAGGGAATGTTA 420

Qy 458 TyrlsThrThrCysLeuValAlaCysAspGluGlyTyrlsArgLeuGluGlySerAspLys 477

Db 421 TATAAGACAAACATGTTTGGTGTGATGAAGGGTACAGACTAGAGGAGGAGTATAG 480

Qy 478 LeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHis 497

Db 481 CTTACTTGTTCAGGAAACACGCTGAGTGGGACAGACCCCGGTGTGTGGAGGCCAC 540

Qy 498 CysSerThrPheGlnMetProLysAspValIleSerProHisAsnCysGlyLysGln 517

Db 541 TGTTCACCTTTTCAGATGCCCAAGATGTTCATCATATCCCCCCACAACTGTGGCAAGCAG 600

Qy 518 ProAlaLysPheGlyThrIleCysTyrlsValSerCysArgGlnGlyPheIleLeuSerGly 537

Db 601 CCAGCCAAATTTGGGACGATCTGCTATGTAAAGTTGGCGCCAAAGGGTTCATTTATCTCGA 660

Qy 538 ValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAla 557

Db 661 GTCAAGAAATGCTGAGATGTACCACTTCTGGAATAATGGATGTGGAGTTCAGGCACCT 720

Qy 558 ValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThr 577

Db 721 GTGTGTAAGACGCTGGAGGCTCTCAAACTCACTGTCTTAAGGACATAGAGGCTAAGACT 780

Qy 578 LeuGluGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSer 597

Db 781 CTGGAACAGCAGAAATCTGCAATGTTTACCTGGCAGATTCACACAGCTTAAGACAACTCT 840

Qy 598 GlyGluLysValSerValHisValHisProAlaPheThrProProTyrlsLeuPheProIle 617

Db 841 GGTGAAAAGGTGTGCTCCGCTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAAT 900

Qy 618 GlyAspValAlaIleValTyrlsThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637

Db 901 GGAGATGTTGCTATCGTATACACGGCAACTGACCTATCCGCAACACAGGCCAGCTGCATT 960

Qy 638 PheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProPro 657

Db 961 TTCCATATCAAGGTTATTGATGCAAGAACCACTCTCTCATAGACTGGTGAGATCTCCACT 1020

Qy 658 ProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAsp 677

Db 1021 CCGCTCCAGGCTTCGGAGAGGTACATGCCGCAAGCTGGGATGAGCCTCAGTTCTCAGAC 1080

Qy 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697

Db 1081 AACTCAGGGGCTGAATTTGGTCAATACCAAGATCATACACAAAGGAGACCTTTTCCCTCAA 1140

Qy 698 GlyGluThrIleValGlnTyrlsThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717

Db 1141 GGGGAGACTATAGTACAGTATACAGCCATGACCCCTCAGGCAATAACAGGACATGTAT 1200

Qy 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737

Db 1201 ATCCATATGTCATAAAGGTTCTCCCTGTGAATCCATTCACACCTGTAAATGGGAT 1260

Qy 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrls 757

Db 1261 TTTATATGCACTCCAGATAATACCTGGAGTCAACTGTATCACTTAACCTTGTGGAGGGCTAT 1320

Qy 758 AspPheThrGluGlySerThrAspLysTyrlsCysAlaTyrlsGluAspGlyValTrpLys 777

Db 1321 GATTTCAAGAGGGTCTACTGACAGATATTTATTTGCTTATGAGATGGGCTCTGGAAA 1380

Qy 778 ProThrTyrlsThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPhe 797

Db 1381 CCAACATATACCACTGAATGGCCAGACTGTGCCAAAAAAGCTTTTTCACAAACCCAGGGTTC 1440

Qy 798 LysSerPheGluMetPheTyrlsAlaAlaArgCysAspAspThrAspLeuMetLysLys 817

Db	1441	AAAGTCCTTTGAGATGTTCTACAAAGCAGCTGTTGTGATGACACAGATCTGATGAGAG	1500	Db	121	GAAGACTGTGTCTGCAGAGAGGGATACAGGCGATCTGGCCAGACCTGTGAACTTGTCCAC	180
Qy	818	PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla	837	Qy	378	CysProAlaLeuLysProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHis	397
Db	1501	TTTTCTGAAGCATTTGGAGAGACCCCTGGGAAAAATGGTCCCATCATTTTGTAGTGATGCA	1560	Db	181	TGCCCTGCCTCGAAGCCCTCCCGAAAAATGGTTACTTTTATCCAAAACACTTTGCACACAC	240
Qy	838	GluAspIleAspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsn	857	Qy	398	PheAsnAlaAlaCysGlyValArgCysHisProGlyLysPheAspLeuValGlySerSerIle	417
Db	1561	GAGGACATTCGCTGACGACTGGAGAGAACCTCGACGCAAAAATATATGCTGGAATATAAT	1620	Db	241	TTCAATGACAGCTGTGGGGTCCGATGTACCCCTGGATTGATCTTGTGGAGAGCAGATC	300
Qy	858	TyrAspTyrGluAsnGlyPheAlaIleGly	867	Qy	418	IleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThr	437
Db	1621	TATGACTATGAAATGGCTTTTCAATTGCT	1650	Db	301	ATCTTATGCTCTACCAATGTTTGGTTCGGCTCAGAGAGCTACTGACAGATGAAGAACA	360
RESULT 13				Qy	438	CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu	457
LOCUS	BD158559	1969 bp	DNA	linear	PAT 17-JAN-2003		
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.						
ACCESSION	BD158559						
VERSION	BD158559.1	GI:27864317					
KEYWORDS	JP 2002191363-A/13402.						
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens						
REFERENCE	Sukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.						
TITLE	Primer for synthesizing full-length cDNA and use thereof						
JOURNAL	Patent: JP 2002191363-A 13402 09-JUL-2002;						
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/13402 PD 09-JUL-2002 PP 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO. PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC 10, C12P21/02, C12Q1/68/C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT CDS (415)..(1653).						
FEATURES	Location/Qualifiers						
source	1..1969 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"						
ORIGIN							
Alignment Scores:							
Pred. No.:	3,55e-192	Length:	1969				
Score:	3073.00	Matches:	547				
Percent Similarity:	99.45%	Conservative:	0				
Best Local Similarity:	99.45%	Mismatches:	3				
Query Match:	30.07%	Indels:	0				
DB:	6	Gaps:	0				
US-09-977-053-6 (1-1842) x BD158559 (1-1969)							
Qy	318	TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIle	337	Qy	678	AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln	697
Db	1	TATGATGCACAGCTTGCCCATTCGGGACATACAAACCTGAACCTCACCAGGAGATC	60	Db	1081	AACTCAGGGGCTGAATTTGGTTCATTACCGAAGATGCATACCAAGGAGACCTTTTCCCTCAA	1140
Qy	338	SerSerCysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerPro	357	Qy	698	GlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp	717
Db	61	AGCAGTTGCATTCCATGTCOGATGAAATCACACCTCTCCACTCGAAGACATCCCTCT	120	Db	1141	GGGGAGCATATAGTACGTATACGACCATGACCTCTCAGGCAATAACAGACATGTGAT	1200
Qy	358	GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis	377	Qy	718	IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp	737


```

Db 1201 ATCCATATTGTCTATAAAAGGTTCTCCCTGTGAAATCCCAATTCACACCTGTAAATGGGGAT 1260
Qy 738 PheIleCysThrProAspThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
Db 1261 TTTATATGCACTCCATATATCTGGAGTCACTGTACATTAATCTGCTGGAGGGCTAT 1320
Qy 758 AspPheThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTdpLys 777
Db 1321 GATTTCACAGAGGCTCTACTGACAGATATATTGTGCTTATGAGATGGGCTCGAAA 1380
Qy 778 ProThrTyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPhe 797
Db 1381 CCAACATATACCACTGAATGCCAGACTGTGCCAAAAAGCTTTTGCAACACCGGGTTC 1440
Qy 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspPheThrAspLeuMetLysLys 817
Db 1441 AAGTCTCTTTGAGATGTTCTACAAAGCAGCTGCTGTGTATGACACAGATCTGATGAAGAG 1500
Qy 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
Db 1501 TTTTCTGAGCATTGGAGACACCTCGGAAAAATGGTCCCATCAATTTTGTAGTATGCA 1560
Qy 838 GluAspIleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsn 857
Db 1561 GAGGACATTCGACTGCAGCTGGAGGAGAACTGACCAAAAAATATTGCTAGATATAAT 1620
Qy 858 TyrAspTyrGluAsnGlyPheAlaIleGly 867
Db 1621 TATGACTATGAAATGGCTTTGCAATGGT 1650

RESULT 14
AK023591 1969 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ13529 fis, clone PLACE1006157, weakly similar
DEFINITION to E-SELECTIN PRECURSOR.
ACCESSION AK023591
VERSION GI:10435568
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Ishibaishi,T., Fujimori,K.,
Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Ishii,S., Kawai,Y.,
Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Negahari,K.,
Masuho,Y. and Kanehori,K.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1969)
AUTHORS Isogai,T. and Otsuki,T.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(8-mail:genomics@ri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source Location/Qualifiers
1. .1969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1006157"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="cloning vector: pMB18SFL3"
415. .1656
CDS

```

```

/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB14617.1"
/db_xref="GI:10435569"
/translation="MLYKTTCLVACDEGYLESGDKLTCCQNSQNDGPEPCVERHCS
TFPMKDVIIISPHNCGKPAKFTICVSCRQGFILSGVREMLKCTTSGKRWNVQAA
VCKDRAPOINCPKDIENAKTLEQDSANVTWQIPTAKDNGSEKYSVRVHPATPPYLF
PIGDVAIVTATDLSGNQASCIPIHKIVDAEPVDMCRSPPPVQVSEKHAASWDEP
TPNDSGAEVLVITRSHQTGDLFQGETIVQVATDPSGNRTCDHIVIKGSPCBIPP
QFVNGDPICTPDNTGVNCTLTCLGVDYDFTSGSDTKYCAVEDGVKMTYTTTWPDKAK
KRFANIGPKSFEMFYKAARCDTDLMKKFSLEALTTLGNKVPSPCSDAEDIDCRLEEN
LTKKYCLEVNYDYENGPAIGN"

Alignment Scores:
Pred. No.: 3,55e-192 Length: 1969
Score: 3073.00 Matches: 547
Percent Similarity: 99.45% Conservative: 0
Best Local Similarity: 99.45% Mismatches: 3
Query Match: 30.07% Indels: 0
DB: 9 Gaps: 0

US-09-977-053-6 (1-1842) x AK023591 (1-1969)
Qy 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIle 337
Db 1 TATGATGACACAGCTTGGCCCATCGGGGACATACAAACCTGAAGCCTCACAGGAGGAATC 60
Qy 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
Db 61 AGCAGTTTGCATTCATGTCGCGATGAAAAATCACACCTCTCCACCTGGAGGACACATC 120
Qy 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
Db 121 GAGACTGTGTCTGCAGAGAGGATACAGGGCATCTGGCCAGACTGTGAACATTGTCCAC 180
Qy 378 CysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHis 397
Db 181 TGGCTGCTGCTGAAGCTCTCCGAAATGGTTACTTTATTCACAAACACTTGCACAAACAC 240
Qy 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIle 417
Db 241 TTCAATGCAAGCTGTGGGGTCCGATGTCACCTCGATTTGATCTGTGGAGGACGATC 300
Qy 418 IleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThr 437
Db 301 ATCTATGCTTACCAATGTTGTGTGCTCGGTTCAGAGACTACTGCAGAGTAGAAGACA 360
Qy 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
Db 361 TGTCTCTCATCTCCGCCAGCGAAACATGGCCACATCAGCTGTTCTACAAAGGAAATGTTA 420
Qy 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
Db 421 TATAAGACACATGTTTGGTTGCTGTGATGAAGGGTACAGACTAGAAAGGACGATGAAG 480
Qy 478 LeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHis 497
Db 481 CTTACTTGTGAGGAAACACAGCAGTGGATGGGCGAGACCCCGGTGTGTGGAGGCCAC 540
Qy 498 CysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGln 517
Db 541 TGTTCACCTTTTCAGATGCCCAAGATGTCAATCATATATATATATATATATATATAT 600
Qy 518 ProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGly 537
Db 601 CCAGCCCAATTTGGGACGATCTGCTATGATGTTGCCCGCCAGGGTTCATTTATCTGGA 660
Qy 538 ValLysGluMetLeuArgCysThrThrSerGlyLysTyrPheAsnValGlyValGlnAla 557
Db 661 GTCAGAGAAATGCTGAGATGTACCACTTCTGGAATAATGGAATGTCGGAGTTCCAGCAG 720
Qy 558 ValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThr 577

```


Db 721 GTGTGTAAGAGCTGGAGGCTCTCAATCAACTGCTCTAGACATAGAGCTTAGACT 780
Qy 578 LeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSer 597
Db 781 CTGGAACACGCAAGATTCCTCCCAATGTACCTGGCAGATTCACACAGCTAAAGACAACTCT 840
Qy 598 GlyGluLysValSerValHisValHisProAlaPheThrProProTyrIleuPheProIle 517
Db 841 GGTGAAAAGGTGTCAATGCGCGTTCATCAGCTTTTCAACCCACCTTACCTTTTCCCAAT 900
Qy 618 GlyAspValAlaIleValThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIle 637
Db 901 GGAGATGTTGATGCTATACAGCGCACTGACCTATCCGGCAACGACGCCAGCTGCAT 960
Qy 638 PheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerProPro 657
Db 961 TTCCATATCAAGTTATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Qy 658 ProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAsp 677
Db 1021 CCCGTCAGGCTCTCGAGAGAGGTATATGCGCGAAGCTGGGATGAGCCTCAGTTCTCAGAC 1080
Qy 678 AsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGln 697
Db 1081 AACTCAGGCGCTGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 698 GlyGluThrIleValGlnIleValThrAlaThrAspProSerGlyAsnAsnArgThrCysAsp 717
Db 1141 GGGGAGATATAGTACATGATATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGAT 1200
Qy 718 IleHisIleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAsp 737
Db 1201 ATCCATATGTCATATAAAGGTTCTCCCTGTGAAATCCCAATTCACACCTGTGAAATGGGGAT 1260
Qy 738 PheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyr 757
Db 1261 TTTATATGCACTCCAGATATATCTGGAGTCACTGATCACTTACTTGTCTGGAGGGCTAT 1320
Qy 758 AspPheThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTrpLys 777
Db 1321 GATTTTCACAGAGGCTCTACTGACAAAGTATTATTGCTTATGAAAGATGGCGCTCGAAA 1380
Qy 778 ProThrTyrThrGluThrProAspCysAlaLysLysValArgPheAlaAsnHisGlyPhe 797
Db 1381 CCAACATATACACATGAAATGGCCAGACTGTGCCAAAACAGTTTTCGCAACACACGGGTTC 1440
Qy 798 LysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLys 817
Db 1441 AAGTCTTTGAGATGTTCTCAAGAGAGCTCTGTTGATGATGATGATGATGATGATGATGAT 1500
Qy 818 PheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAla 837
Db 1501 TTTTCTGAAGCATTTGGAGACGACCTCGGAAAATGTTCCCATCATTTTGTAGTATGCA 1560
Qy 838 GluAspIleAspCysArgLeuGluGlnLeuThrLysLysTyrCysLeuGluTyrAsn 857
Db 1561 GAGGACATTTGACTGAGACTGGAGAGAACCTTGACCAAAAATATTTGCTAGATATATAT 1620
Qy 858 TyrAspTyrGluAsnGlyPheAlaIleGly 867
Db 1621 TATGATATGAAATGGCTTTCGAAATTTGAT 1650

RESULT 15
LOCUS BD229966 1760 bp DNA linear PAT 17-JUL-2003
DEFINITION Method and composition related to EGF repeat-containing polypeptide.
ACCESSION BD229966
VERSION BD229966.1 GI:33039736
KEYWORDS JP 2002526077-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

OS Homo sapiens (human)
PN JP 2002526077-A/6
PD 20-AUG-2002
PF 20-SEP-1999 JP 2000574256
PR 18-SEP-1998 US 09/157308
PI JOHN E FORD, JULIO J MULERO, GEORGE YEUNG, JULIE MONTGOMERY PC
C12N15/09 A61K38/00, A61K48/00, A61P1/00, A61P1/02, A61P3/10 PC
, A61P7/00, A61P7/02, A61P9/00, A61P11/00, A61P11/06, A61P17/00, PC
A61P17/02,
PC A61P19/00, A61P19/02, A61P19/10, A61P21/00, A61P21/04, A61P25/00,
PC A61P25/00,
PC A61P25/14, A61P25/16, A61P27/02, A61P29/00, A61P29/00, A61P31/00,
PC A61P31/04, A61P31/12, A61P31/16, A61P31/18, A61P31/22, A61P33/00,
PC A61P35/00,
PC A61P35/02, A61P37/06, A61P43/00, C07K14/485, C07K16/22, C12N1/15,
PC C12N1/19,
PC C12N1/21, C12N5/10, C12Q1/68, C12N15/00, C12N5/00, A61K37/02 CC
w= a or t

PH Key Location/Qualifiers
FT CDS (4) . (1407) .

FEATURES
source

1. 1760
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,36e-146 Length: 1760
Score: 2374.00 Matches: 433
Percent Similarity: 96.07% Conservative: 7
Best Local Similarity: 94.54% Mismatches: 17
Query Match: 23.23% Indels: 1
DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x BD229966 (1-1760)

Qy 845 GluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPhe 864
Db 3 GAGGAGAGCGCCGACCAATCTCTGACTGACTCTTACTTACGACTATCCAGCTGGCTGG 62
Qy 865 AlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPhe 884
Db 63 GCGATGGGACCGAGGGGCTGGGGCGGAGCTCATAGGACTCTTACGATGACTTG 122
Qy 885 LeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLys 904
Db 123 CTGGAGCTGAGCAGAGAAACACGCCACCATCGCAATGCCAAGTACTCACCGAGCTAT 182
Qy 905 ArgSerAla-ProLeuSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerVa 924
Db 183 AAGAAGCGACGACCATATCTGACTATAAAATTAAGTTAAATTTTAAACATCACAGCTAGTGT 242
Qy 924 LProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnArgLeuLeuGlu 944
Db 243 GCGATTTACCGATGAAAGAAATGATACCTCTGAATGGGAAATCAGACAGCATCTCTTCA 302
Qy 944 nThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSe 964
Db 303 GACATTTGAACTATCAACAATAAAGGACTCTCAACAAAGACCCCATGTATTC 362
Qy 964 rPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAl 984
Db 363 CTTTCAGCTTGCATGAGAAATATCTTATAGCCGACAGCAATTCATTAGAAACCAAAAGGC 422

QY 984 aSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLe 1004
 DB 423 TTCCCCCTTCTGCAGACACAGGCTCAGTGTGAGAGGGCGTATGTGTCAATTTGCCCTT 482
 QY 1004 uGlyThrTyrrAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrrGl 1024
 DB 483 GGGAACTTATTATATCTGGACATTTACCTGTGAAGCTGCGGATCGGATCTATCA 542
 QY 1024 nAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyrrThrGluTyrrIl 1044
 DB 543 AGATGAGAGAGGCAACTTGTAGTCAAGCTTTGCCCTCTGGGATGTATACAGGAATATAT 602
 QY 1044 eHisSerArgenIleSerAspCysValalaglnCysLysGlnGlyThrTyrrSerTyrrSe 1064
 DB 603 CCAATCAAGAAACATCTCTGATGTGAAGCTCAGTGTAAACAGGACCCCTACTCATACAG 662
 QY 1064 rGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrrGlnProLysPheGlySerAr 1084
 DB 663 TGGACTTGAGACTTGTGATCTGTCCACTGGGCACCTTATCAGCCAAATTTGGTTCCCG 722
 QY 1084 gSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSe 1104
 DB 723 GAGCTGCTCTGCTGTCCAGAAACACCTCAACTGTGAAAGAGGAGCGGTGAACATTTTC 782
 QY 1104 rAlaCysGlyValProCysProGluCysPheSerArgSerGlyLeuMetProCysEhi 1124
 DB 783 TGCATGTGGAGTTCCTTGTGCCAGAGGAAATTCCTGCGGTTCCTGGGTAAATGCCCTGTCA 842
 QY 1124 sProCysProArgAspTyrrTyrrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysPr 1144
 DB 843 CCCATGCTCTGTGACTATTACCACTTAATCAGGGAGGCTTCTGCTGGCTGTCTC 902
 QY 1144 oPheTyrrGlyThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSe 1164
 DB 903 CTTTATGGAATACCCCATTCCTGCTGTTCAGATCCATCAGAGATGTTCAGTTTATAG 962
 QY 1164 rSerThrPheSerAlaAlaGluGluSerValProAlaSerLeuGlyHisIlely 1184
 DB 963 TTCACATTTCTCAGCGGAGAGGAAAGTGTGTGCCCCCTGCTCTCTTGGACATATTAA 1022
 QY 1184 sLysArgHisGluIleSerSerGlnValPheHisGluCysPheAsnProCysHisAs 1204
 DB 1023 AAAGAGGATGAAATCAGCAGTCAGGTTTCCATGATCTTCTTTAAACCTTGCACAA 1082
 QY 1204 nSerGlyThrCysGlnGlnLeuGlyArgGlyTyrrValCysLeuCysProLeuGlyTyrrTh 1224
 DB 1083 TAGTGGAACTGCCAGCACTTGGCGTGGTATGTCTGTCTCTGTCCACTTGGATATAC 1142
 QY 1224 rGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAengl 1244
 DB 1143 AGGCTTAAAGTGTGAACAGACATCGATGAGTGACGCCACCTGCTTGCCTCAACANTGG 1202
 QY 1244 vValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrrThrGlyGl 1264
 DB 1203 AGTTTGTAAAGACCTAGTTGGGGAATTCATTTGTGAGTCCCATCAGGTTACAGGTCA 1262
 QY 1264 nArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIleCy 1284
 DB 1263 GCGGTGTGAAGAAATATATAATGAGTGTAGTCCAGTCTCTTGTGTAAATAAAGGAATCTG 1322
 QY 1284 vValAspGlyValAlaGlyTyrrArgCysThrCysValLysGlyPheValGly 1301
 DB 1323 TGTGTGTGTGTGGCTATCTGTTCACATGTGTGAAGGAATTTGTAGGT 1374
 RESULT 16
 LOCUS AX870332
 DEFINITION Sequence 5237 from Patent EP1074617.
 ACCESSION AX870332
 VERSION AX870332.1 GI:40025195
 KEYWORDS Homo sapiens (human)
 SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J.,
 Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 TITLE Primers for synthesising full-length cDNA and their use
 JOURNAL Patent: EP 1074617-A 5237 07-PEB-2001;
 Research Association for Biotechnology (JP)

FEATURES
 source
 1..765
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 5,4e-86 Length: 765
 Score: 1447.00 Matches: 251
 Percent Similarity: 98.82% Conservative: 0
 Best Local Similarity: 98.82% Mismatches: 3
 Query Match: 14.16% Indels: 0
 DB: Gaps: 0

US-09-977-053-6 (1-1842) x AX870332 (1-765)

QY 195 ThrAspGlyTyrrSerAsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSer 214
 DB 2 ACTGATGGAATATTCCTAATGGGGAGACCTTAGACCATTGCGAGCTGCTGCGAGATTCA 61
 QY 215 GlyValGluIlePheThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMet 234
 DB 62 GGAATGGAGATCTCTCACTTTGGCATATGGCAAGGGAACATTCGAGAGCTGATGACATG 121
 QY 235 AlaSerThrProLysGluGluHisCysTyrrLeuLeuHisSerPheGluGluPheGluAla 254
 DB 122 GCTTCCACCCTCAAGAGGAGGAGCCTGTACTCTCTACACAGTTTGTGAAGAAATTTAGGCT 181
 QY 255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMet 274
 DB 182 TTAGCTCGCGGGCATTTGCATGAAGATCTACCTTCTGGGAGTTTATTTCAAGATGATATG 241
 QY 275 ValHisCysSerTyrrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCys 294
 DB 242 GTCCACTGCTCATATCTTTTGTATGAAGGCAAGCACTGCTGTGACCGAATGGGAAGCTGC 301
 QY 295 LysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrrTyrrGlyLys 314
 DB 302 AAATGTGGGACACACACAGGCCATTTTGTGTGATCTGTGAAAAGGGGTATTTACGGGAAA 361
 QY 315 GlyLeuGlnTyrrGluCysThrAlaCysProSerGlyThrTyrrLysProGluGlySerPro 334
 DB 362 GGTCTCAGTATGAATGCAAGCTTCCCATCGGGACATACAACTGAGGCTGACCA 421
 QY 335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProGlySer 354
 DB 422 GGAGGAATCAGCAGTTGCATCTCCATGTATGATAAATCACACCTCTCCACCTGGAAGC 481
 QY 355 ThrSerProGluAspCysValCysArgGluGlyTyrrArgAlaSerGlyGlnThrCysGlu 374
 DB 482 ACATCCCTCTGAGACTGTGTCTGACAGAGAGGGATACAGGGCATCTGGCCAGACCTGTGAA 541
 QY 375 LeuValHisCysProAlaLeuLysProGluAsnGlyTyrrPheIleGlnAsnThrCys 394
 DB 542 CTTCGCCACTGCTCTCCCTGAAAGCTTCCGAAATGTGTTACTTTATCCAAACACTTGC 601
 QY 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
 DB 602 AACACCACTTCAATTCAGCTGCTGGGGTCCGATGTCACTTGTGATTTGATCTTGTGGGA 661
 QY 415 SerSerIleLeuLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTyrrCysArg 434
 DB 662 AGCAGCATCATCTTATGTCTTACCCTTGTGGTTCGGTTCGGTTCGGTTCGGTTCGGTTCGGT 721

```

QY 435 ValArgThrCysProHisLeuArgGlnProtyrHisGlyHis 448
DB 722 GTAAGAACATGCTCTATCTTTGTGATGAAGCAAGGACTGCTGTGACCGAATGGCAAGCTGC 301

RESULT 17
BD150394 765 bp DNA linear PAT 17-JAN-2003
LOCUS Primer for synthesizing full-length cDNA and use thereof.
DEFINITION BD150394
ACCESSION BD150394
VERSION BD150394.1 GI:27856152
KEYWORDS JP 2002191363-A/5237.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota.T., Isegai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J.,
Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 5237 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/5237
PD 09-JUL-2002
PP 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof PH Key
Location/Qualifiers
FT source 1..765
FT /organism='Homo sapiens (human)'.
FEATURES
source
Location/Qualifiers
1..765
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 5.4e-86 Length: 765
Score: 1447.00 Matches: 251
Percent Similarity: 98.82% Conservative: 0
Best Local Similarity: 98.82% Mismatches: 3
Query Match: 14.16% Indels: 0
DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x BD150394 (1-765)

QY 195 ThrAspGlyTyrSerAsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSer 214
DB 2 ACTGATGGATATCCAAATGGGGAGACCCCTAGACCAATTCGACGCTGCAGATTCA 61

QY 215 GlyValGluIlePheThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMet 234
DB 62 GGAGTGGAGATCTTCATCTTTGGCATATGCGAAGGAGACATTCGAGAGCTGAATGACATG 121

QY 235 AlaSerThrProGlyGluHisCysTyrLeuLeuHisSerPheGluIleGluAla 254
DB 122 GCTTCACCCCAAGGAGGAGCACTGTACCTGCTACACAGTTTGAAGAAATTCAGGCT 181

QY 255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMet 274
DB 182 TTAGCTCGCCGGGCATTCGATGAGATCTACCTTCGGGAGTTTATTTCAGATGATATG 241

QY 275 ValHisCysSerTyrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCys 294
DB 1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLys 1281

```

62 ACAGGTGAGCGGTGTCGAGAAATATATAATGAGTGTAGCTCCAGTCTCTGTTAAATATA 121
 1282 GlyValCysValAspGlyValAlaGlyTyrArgCysThrCysValysGlyPheValGly 1301
 122 GGAATCTGTGTGATGTTGGTGGCTATGCTGTGCACATGTGTGAAGATTTGTAGGC 181
 1302. LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
 182 CTGCATTGTGAAACAGAGTCAATGATGTCAGTCAACCCATGCTTAATATGAGTGC 241
 1322 CysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArg 1341
 242 TGTGAAGACAGGTTGGGGATTTCTGTGCAATGCCACCTGATTTTGGGTACCCGA 301
 1342 CysGlyValAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLys 1361
 302 TGTGGAAGAACCTGCATGAGTCTCTCAGTCAACCCATGCAAAATGAGGATCTACTGTAA 361
 1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGlu 1381
 362 GACGGTGCCTAATGATCTCAGTCTCTGTGTGAGCTGGCTTTCAGGATCAGTCTGAA 421
 1382 LeuAsnAlaAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
 422 TTCAACATCAATGAATGTGAGTCTAATCCATGATGAGAAATCAGGCCACCTGTGTGATGA 481
 1402 LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
 482 TTAATATCAATACAGTTGTAAATCTCAGCCAGGANTTTTCAGGCAAAAGGTGTGAACAGAA 541
 1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
 542 CAGTCTACAGGCTTTAACTGATTTTGAATTTCTGGCATCTATGATATGTCATGCTA 601
 1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSerAsp 1461
 602 GATGGCATGCTCCATCTCTCCATGCTCTAACTGATCTAACTGATCTCTGATGAATCTCTGAC 661
 1462 AsnMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGly 1476
 662 GACATGAATATGGAACCAATCTCTATGAGTGTGATAACGGC 706

RESULT 19
 AC018297 18583 bp DNA linear HTG 09-DEC-1999
 LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
 DEFINITION AC018297
 ACCESSION AC018297.1 GI:6552894
 VERSION HTG: HTGS PHASE2.
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE Adams, M. and Venter, J.C.
 AUTHORS 1 (bases 1 to 18583)
 TITLES Direct Submission
 JOURNAL Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 COMMENT This sequence was identified as CDM:10214228 by the submitter.
 For more information on this record e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 Location/Qualifiers
 1..18583
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 ORIGIN
 Alignment Scores:

Pred. No.: 2,82e-76 Length: 18583
 Score: 1328.00 Matches: 518
 Percent Similarity: 32.44% Conservative: 256
 Best Local Similarity: 21.71% Mismatches: 718
 Query Match: 13.00% Indels: 900
 DB: 2 Gaps: 88
 US-09-977-053-6 (1-1842) x AC018297 (1-18583)
 QY 6 AlapheCysCysTrpGlyLeuAlaLeuVal-----SerGlyTrpAla 19
 DB 8021 GCACAAATGCTGTGGTAAAGTCTTATTAATTCAGATTACATAGTTCAATACATAATCT 8080
 QY 20 ThrPheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGlyThrAla 39
 DB 8081 TCAATCTTGAATCTTCTAGATGTGTGGCCACATCTCTGGCCGATGTC-CCGAGCTGCA 8139
 QY 40 ProGlyAlaProGlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySer 59
 DB 8140 CCGGGA-----ATTGGAGCTTGGCCTCCAGTCCGACACCGCTTTTCGGTACG 8187
 QY 60 ArgValGlu-----ArgLeuGlyGlnAlaPheArgArg-----ArgValArgLeuLeu 75
 DB 8188 CTGGTCAGCTTTTACATGTCCTCATTTGGACAGAGTTTCCACCAGCCAGACCGCTGGT 8247
 QY 76 ArgGluLeuSerGluArgLeuGluLeuValPheLeuValAspAspSerSerValGly 95
 DB 8248 ACCGATG-TCTGGCGGTGGCAACTGGAGTGTCTCTACATACCCAAAGTGTCTAG- 8304
 QY 96 GluValAsnPhe-----ArgValAsnPhe----- 99
 DB 8305 GAGTGTGCCTTACCTATTTCCATAATTTGATATCCATAATTCATATGCAATCCCTTCCA 8364
 QY 100 ---ArgSerGluLeuMetPheValArgGlyLeuLeuSerAspPheProValValProThr 118
 DB 8365 CAGAGGTCTACTCGGTCTCTGTGCCAATAATCGCAACGGTTTCTCATTTGCTCTCGA 8424
 QY 119 AlaThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAsp 138
 DB 8425 ACCTAACCTATCGGTATAGCAATGTACAGTCTAGCGCGGTCTTGCCTTCGCTCGG 8484
 QY 139 TyrIleSerThrArgArgAla-----ArgGlnHisLys 149
 DB 8485 GTGCTCGATCGAGAGATC-TGCTGTCTCGGATGGCGCTTGGAGCGACAGCCCCAC 8543
 QY 150 Cys---AlaLeuLeuGlnGlnIleProAlaIleSerTyr----- 162
 DB 8544 TGCATGCGCTCCCGAGTGTGCGAGCGCTGCCGGAAGTGGCACAGCCCAACGTCACTGCTG 8603
 QY 163 ArgGlyGlyGlyThrTyrThrLysGlyValaPheGlnGlnAlaAlaGlnIleLeuLeuHis 182
 DB 8604 AATGGAGTGGT----- 8615
 QY 183 AlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGly 202
 DB 8616 ---CGCAGCTACGGCACCATTTGTCCAGTGTGAGTGTGAGCGGGCTACGAGCGCAATGCG 8672
 QY 203 AspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGly 222
 DB 8673 CATCCC-----GTGCTGACCTGTATGTCCAGCGC 8702
 QY 223 IleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHis 242
 DB 8703 ACCTGGAGTGTGTATGTATCCCAAGATGCAAG-----CGCAAGCGG 8741
 QY 243 CysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAla----- 256
 DB 8742 TGCTTC-----GAATCCCGACCATTTGCCACGGCTTTGTGTGGAC 8783
 QY 257 -----ArgArgAlaLeuHisGluAsp-----LeuProSerGlySer 268
 DB 8784 TCGACGCGAGCCTACCTCTTTCGGCGATGAGCCAGGGTGCAGTGTCTCAAGGGCTACAAA 8843

QY 269 PheIleGlnAspMetValHisCysSerTyr-----Leu 280
 DB 8844 CTGATCGGAGCAACATCATGCGCTGCAGCGAGGCCCGAGAGTTGAGAGCGCCGCCGAGC 8903
 QY 281 CysAspGluGlyLysAspCys-----CysAspArgMetGlySerCysLysCys 296
 DB 8904 TGGAGGACATCAACGAGTGCAGCTCTCTCGAGTGGAC---CTAACCAACCGAGTGC 8960
 QY 297 GlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly----- 313
 DB 8961 CAGAACACGACCGCTCTCTTCCACTGCCAGTGCAGGCGGATTCAC-GGCTACCACCGA 9019
 QY 313 ----- 313
 DB 9020 GTGTCGGCCCGTCGTGTGATTTGGCTGTGGGTAAATCGAGGCATACCGGATGACAGCATCAC 9079
 QY 314 -----LysGlyLeuGlnTyrGluCysThrAla----- 322
 DB 9080 CACCTCGGTCAGTGAGCCGGCTACAGCAAGGAGCAGCTGGCTTGAACACGAATGGCTG 9139
 QY 323 -----CysProSerGlyThrTyrLysProGluGlySer----- 333
 DB 9140 GTGCGGTGGCTCTTCGAGCGCTGTGTGCCAACCTGTGATCTCATCGACCTGAAGGCACCCAC 9199
 QY 333 ----- 333
 DB 9200 CATTCGTGCGTTCCTCCGACCATGTCTCGTCGAGCGTCCCGATGGCAATGTGGCTTCAG 9259
 QY 333 ----- 333
 DB 9260 CTCGGGGTGGCTGTCAGTACACCAACGATCTGACGGATGTGTTCAAGGATTATGCCAA 9319
 QY 333 ----- 333
 DB 9320 TCCCGAGCGCACTGCGCTGCAATTCGGCATCTCGGAGCCACGCTCTCCATCTTAAACCT 9379
 QY 334 -----ProGly----- 335
 DB 9380 GCCCTGCGCATCGAAGCTGCTATATTGCTTCGCGATCCAGGACTACGTGGGTGGCCG 9439
 QY 336 -----Gly-IleSerSerCys-----IleProCysProAspGluAsnHis 349
 DB 9440 CTGCTCGCGCATGGAGCTGATGGCTGCAGCGCTTGGATTGGTGGACATCAACGAGTG 9499
 QY 349 rSerProGlySerThrSerProGluAspCysVal----- 361
 DB 9500 CAGCAAGAACATGGCGCTGTGACCAAGATGTCATCAACTCACCGGCGGATTTGCCCTG 9559
 QY 362 -----CysArgGluGlyTyrArg----- 367
 DB 9560 TGGCTGCACACTGGCTACCACTGTATACCTTCCAAACGGCAGCGGTGCTATCATATCA 9619
 QY 368 -----AlaSerGlyGln-----ThrCysGluLeu 375
 DB 9620 AGCTCCGATCCGGGAGCGTGTGATGGTGCACACCTATCAGCGCAACAGACCTGTGTTC 9679
 QY 375 uValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnHisCysAs 395
 DB 9680 TCTCATGTGTCGGAACCTGGAGCGCCCGAGAAATGTTCAACTCTCTGAGCGACAGAACGA 9739
 QY 395 nAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySe 415
 DB 9740 CTATCATTTTGGCGATGTGGCTTCGATGGCCACTTTGGCTATCATGTAGCGGGCAG 9799
 QY 415 rSerIleIleLeuCysLeuProAsnGlyLeuTyrSerGlySer-----GluSerTyrC 433
 DB 9800 CTGGCGGCGCTGTGCTCTCTCAGCGGTCAGTGGAAACGCGCAGCGTACCGGAGTGCATTTG 9859
 QY 433 sArgValArgThr----- 437
 DB 9860 TAAAGTTTATACATTCTCAGATTCTCAATGAATTCCTTAAACAAACGAAATGCTCAC 9919
 QY 437 ----- 437

DB 9920 TTTTCAGATGCCAATGCGTTTCCCTGCCCGCATGACAAGTTGAGGGTCTGACTGTGGCC 9979
 QY 438 -----CysProHisLeuArgGln----- 443
 DB 9980 CGCCCCGATCCCGAATCGTTCTAGTGGCCCTTCCGTTGACAATGTGACCAATTAGTCGGGA 10039
 QY 443 ----- 443
 DB 10040 TGGCCGGGACGCCAACTGAGAGCCACCGCTTCTCTGTTTCCGGCAGTGGCTGTAGCAT 10099
 QY 444 -----ProLysHisGlyHisIleSerCys----- 451
 DB 10100 CCCAAGCCCGCTCTGCCCGATTACTGGCTATCCGGAATGCAGCCCTCTTGTTCGCCGAGTG 10159
 QY 452 -----SerThrArgGluMetLe----- 457
 DB 10160 GATTGCTACTCACCCATGCCAACGCCCGCGGAGAAATACGACAGCTTTGTG-GACACTCG 10218
 QY 457 uTyrLysThrThrCysLeuValAlaCysAspGluGlyTyrAsnLeuGluGly----- 474
 DB 10219 CTATCAGAGCAGCTTCTTCTTGGCTGCCAGAACACCTTTAAGTTGGCTGGACAGCGGG 10278
 QY 475 -----SerAspLysLeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProAr 492
 DB 10279 TCGTCACGACATGTGTTGTTGGAGCCGATGTATCTGGACTTTGGAGATCTTCG 10338
 QY 492 gCysValGluArgHisCysSerThrPheGlnMetProLysAsp----- 506
 DB 10339 CTGTGAGGACCTGTGTGCGAGGATCCGGAGACCGGAGATGTGTGCGCAGATTGCAAG 10398
 QY 507 -----ValIleIle 509
 DB 10399 CAGCTATGAGCAGAGCTCGAGGTGTACTTCGGCTGCAATCGTCTGGCTGATCTCTGAT 10458
 QY 509 sSerProHisAsn----- 513
 DB 10459 CAATCGGAGCCCATTTACATGATACGAGCCAGAGTCAAGGTCAACAGCCTTTGGG 10518
 QY 513 ----- 513
 DB 10519 ATTAAGTTCCGCGAGGNTTCGGATTGCGCCATCAATGCCACTCGAGCGACCCCAATTA 10578
 QY 514 -----CysGlyLysGln----- 517
 DB 10579 CGAGGCCAAGAACATCCGTCTCACTCGGCCACTGGCTGTGTGCAACAGGAGGCCCTT 10638
 QY 517 ----- 517
 DB 10639 CACCTATGTGAGCGTGGATCTGGGTGATGATCTATCGAGTCAAGGCGATTCTGTGTGAAGG 10698
 QY 517 ----- 517
 DB 10699 TGTGGTTACCAACGACATGTGTGGCGAGCCCGAGAGATTCGGTTCTTCTTACAAACAGC 10758
 QY 518 -----ProAl 519
 DB 10759 TGAGAGCGAGAACTAGTGGTGTACTTCTCCCAATTTCAATCTGACCATCGAGATCCAGG 10818
 QY 519 aLysPheGlyThrIleCysTyrValSer-----CysArgGlnGlyPheIleLe 535
 DB 10819 CAACTACGCGAGCTGGCCATGATCAGCTGCCCAAGTTGCTGTCAGGCTCGCTTGTGTGAT 10878
 QY 535 uSerGlyValLys-----GluMetLeuArgCysThr 545
 DB 10879 CCTTGAATAGTGAAGTACATGACACCGCTGTCTGAAGTTGAGATTGATGGCTGGCGGA 10938
 QY 545 rThrSerGlyLys-----TriAsnValGlyValGlnAlaValCysLys 560
 DB 10939 GGAGCGGAACAGGAAACCACTCTCTGGCTACGATACGAGTAC-----TCCCGGTGGT 10992
 QY 560 sAspValGluAlaProGlnIle-----AsnCysProLysAsp-----IleGluAlaLysThrLe 578
 QY 578 ----- 578

Db 10993 GGACAGACACACCCATCTTCCAAAACCTGCGCGACGACCAACCAATCTGTGTGCGACGCGA 11052
Qy 578 uGluGlnInAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGI 598
Db 11053 TGAGAAATGGAGGAGTACTACCGGTAACTTCCAGCAACCCACGCGGTGGACAACTCCGG 11112
Qy 598 yGluLysValSerValHisValHisProAla-----PheThrProProTyrIleuPhePr 616
Db 11113 ATCGATTGCCCGCTGGAGATCAAGCCACAGAACTTCCGACACCCAGCTACATTTTCAA 11172
Qy 616 olleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCy 636
Db 11173 G-----GATACGGTTGTAAAGTACGTGGCGCTTGTACGATGGCAATGTGGCCATCTG 11226
Qy 636 sIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerPr 656
Db 11227 CGAGATCAACATCAGCGTGCCGATGTAACACCACTGCTGCAG---TGC----- 11275
Qy 656 oProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSe 676
Db 11276 -CCCGAGAGCTATGTGATTGAGCTAGTGGATCGCGAGACAGCTACACTGTGAACITCAA 11334
Qy 676 rAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu----- 694
Db 11335 CGATACCCGGAAGG-----ATCAAGACCTCCGACGACACAGAGATGTGAGTT 11385
Qy 695 -----PheProGlnGlyGluThrIle-----ValGlnTyrTh 705
Db 11386 GCAGTTTCAGCCCGAGAGTGCACATCAAGATCGGAACCTTCGAGAACGTGACCGTCAC 11445
Qy 705 rAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySe 725
Db 11446 GGCAACGGATAAGTACAAACACCGCGCGCTGCCACTTCAGGTCTCTGTGAAGGCTTC 11505
Qy 725 rProCys-----GluIleProPheThrProValAsnGlyAspPheIleCysThrPr 742
Db 11506 ACCCTGGTGGAGCTGGGAGCTCCAG---CCGCGCGGAATGGTGCCATCAATGCCCTGCC 11562
Qy 742 oAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGI 762
Db 11563 TGGTGATCGTGGTATCGAATGCAATGTCACGTGCAGCCAGGATTCGGTTTCACCGAGG 11622
Qy 762 ySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyrThrTh 782
Db 11623 CGAAACCACTGAAGACCTTCTCTCGCGAGACATCATCGTCTGTGGGCTCCACGTCCTGGGT 11682
Qy 782 rGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLys---SerPheGI 801
Db 11683 G-----CCCGACTGCTGTGCGAGACACGAGACGCGCGCTTACCACGTGACCGCTC 11736
Qy 801 uMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAl 821
Db 11737 CATTAACCTACCGCGCAATGAGCAGTGGGCCAATCTCTGTGTGTCAGTACCAGGAGGT 11796
Qy 821 aPheGluThrThrLeuGly-----LysMetValProSerPheCysSerAsp----- 836
Db 11797 GCTGGCACACCACTATGGCGAGCTCAACACAGTGTCTCTCGAGCGCTGCTCCGCGGTGAA 11856
Qy 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuTh 849
Db 11857 TGTCAACATGAATGTGACCTTTGTGAAGTCTGTGCCCATCTGCTGGAGGAGAATGTGT 11916
Qy 849 rLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGI 869
Db 11917 CAAGATG----- 11923
Qy 869 yGlyTyrGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGI 889
Db 11924 -----GACTTCATCTCTCCAT-- 11941
Qy 889 nGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLe 909
Db 11942 -----CTGCCGCTGTGGTGCAGCGCGAGCTGTACGACCT 11976

Qy 909 uSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspGI 929
Db 11977 GTGGCGCTCCACGCTGAACCTGATCTTTGATCTG-----AGTGTACCTATGCCAGTGC 12030
Qy 929 u---ArgAsnAspThrLeuGluTyrGluAsnGlnArgLeuLeuGlnThrLeuGluTh 948
Db 12031 CGTGATCGATGACCTTTTGAACATTCGTAACATCGTAACAGTGTCTCCGCTACGCGC 12090
Qy 948 rIleThrAsnLysLysArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAl 968
Db 12091 CCTCAAGTCGCAATCTCGGAGGATTTAACTGCAAT-----GT 12129
Qy 968 aSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerProPheCy 988
Db 12130 GGGCGAGTACTGAACATGACACCGCATGTCCG----- 12166
Qy 988 sArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTy 1008
Db 12167 -----CGTTGCTGCACCTGTCCCGCGCAACGTATGT 12198
Qy 1008 rAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGluGI 1028
Db 12199 GTCAGAGGGTCAAGAACAGCTGCACTACTGCCCGAGGGGCTACTACCAGAACCGGTGACCG 12258
Qy 1028 yGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAs 1048
Db 12259 CAGGGAACTGCTGCGTGGCGCGCGGAACTTACACCAAGAGGAGGCGGCCACCATGTC 12318
Qy 1048 nIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluTh 1068
Db 12319 GCAGCGAGTGCATTCCTCCGCTGGGTATGGCACTACTTACCCACCGGACTGTGTGCC 12378
Qy 1068 rCysGluSerCysProLeuGlyThrTyr-----GlnProLysPheGlySerArgSe 1085
Db 12379 GTGCTCGAGTGTCCGGTAACTCATTCCTGCGGAACCAACCAACCGGTGATTCAGGA 12438
Qy 1085 rCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAl 1105
Db 12439 TTGCAGCGCTGTGCGGACAGAGCTTCACTACAGCGCGCTGCTCGAACAGGATCT 12498
Qy 1105 aCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisPr 1125
Db 12499 GTGTGGCGCCAGTGTGCGCGCGGAGAGTACTTCCGCCACCGGACTGGCACCTGTCTGCC 12558
Qy 1125 oCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPh 1145
Db 12559 CTGCCCACTGCATCATTTACCAGGAGCGCGGGTGGCAGAGCTGCACACGAGTGTCCGAG 12618
Qy 1145 eTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSerSe 1165
Db 12619 TAACATGAGAACCTTTCACCGCTCAAGGGACGCGCAACAGTGC----- 12664
Qy 1165 rThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGlyHisIleLysLy 1185
Db 12665 -----AAGCGGTGGTA----- 12676
Qy 1185 sArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisAsnSe 1205
Db 12677 -----TGTGGTGAAGGTGCTTCCACGACGCG 12702
Qy 1205 rGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGI 1225
Db 12703 CGGACTGTGTGTCGCCATGGCCATGATCCAGTGTCTTCTGTCCGCGCGGATTTCTCGG 12762
Qy 1225 yLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGlyVa 1245
Db 12763 ACCTGCTGTGGAAACAGACATTCGACGAGTGGCGCTCCAGCGCTGTCTACATATGGTGTCA 12822
Qy 1245 lCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGluAr 1265
Db 12823 GTGCAGAGTATTCGCGCAGGGCTATCGCTGTGAGTGGCGGCTGTGATCTCGGCGATCA 12882

[illegible]

Qy	333	-----	333	-----	333
Db	32518	CTCGCGGTGCGTCTGCAGTACACCAACGATCTGACGGATGTGTTCAAGGATTATGCCAA	32459		
Qy	333	-----	333	-----	333
Db	32458	TCCCGACGGCAGTCGCGTCGAATTCGCGATCTCGAGGCCAGCTCTCCATCTCTAAACCT	32399		
Qy	334	-----	335	-----	335
Db	32398	GCCCTGCCCATCGAAGCTCGCTATATTTCGCTTCGCCATCCAGGACTACGTGGGTGGCC	32339		
Qy	336	-----	349	-----	349
Db	32338	CTGTCTCGCATGGAGCTGATGGGCTGCACGGCGCTTGGATTGCTGGACATCAACAGCTG	32279		
Qy	349	rSerProProGlySerThrSerProGluAspCysVal	361	-----	361
Db	32278	CAGCAAGAACAATGGCGGCTGTGACACAGAAGTGCATCAACTCACCGGCGGATTTGCCTG	32219		
Qy	362	-----	367	-----	367
Db	32218	TGCGTGCACACTGGCTATACAGCTGTACACCTCCAACGGCAGGCTGCATCATCGA	32159		
Qy	368	-----	375	-----	375
Db	32158	ACGCTCCGAATCCGCGCAAGCTGATGGTGACACCTATCAGCGCAACAAGACTGTGTTC	32099		
Qy	375	uValHisCysProAlaLeuIysProProGluAsnGlyTyrPheIleGlnAsnThrCysAs	395		
Db	32098	TCTCATGTGTCCCGAACTGGAGCGCCGAGAAATGTCACTCTCCGACGCAAGAACGA	32039		
Qy	395	nAsnHisPheAsnAlaCysGlyValArgCysHisProGlyPheAspLeuValGlyse	415		
Db	32038	CTATCACTTTGGCGATGTGTGCGCTCTCCAGTGCCACTTTGCTACATCATCAGCGGCAG	31979		
Qy	415	rSerIleLeuLeuLeuProAsnGlyLeuTrpSerGlyser	433		
Db	31978	CTCGCGGCGCTGTGCGCTCTCCAGCGGTCACTGGAACGCCAGCTACCGGAGTGCATTG	31919		
Qy	433	sArgValArgThr	437	-----	437
Db	31918	TAAGTGTATACATTACAGATCTCAATGAATTCCTCAATTAACAACGAAATGGTCAC	31859		
Qy	437	-----	437	-----	437
Db	31858	TTTTCAGATGCCAAATGCGTTTCCTCGCCCGATGACAGTGTGAGCGTCTCGACTGTGCC	31799		
Qy	438	-----	443	-----	443
Db	31798	CGCCCGATCCCGAATCCGTTCTAGTGCCCTTCCTCGTGAATGCAATGACCAATACGTGGGA	31739		
Qy	443	-----	443	-----	443
Db	31738	TCGCGGGACGCCAATGAGACGCCACCGTCTCTCTGTTTCGCGCAGTGGGTACGAT	31679		
Qy	444	-----	451	-----	451
Db	31678	CCCAAGCCCGGTCTGCCCGATTACTGGCTATCGGAAATGCAACCCCTCTGTCCCGCAGTG	31619		
Qy	452	-----	457	-----	457
Db	31618	GATTGCTACTCACCCATGCCAAGCCCGCGCAGAAATACGACAGATTGTG-GACACTCG	31560		
Qy	457	uTyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly	474	-----	474
Db	31559	CTATCAGACGAGCTCTCTCTTTGGCTGCCAGAACACCTTTAAGTTGGCTGGACAGCGG	31500		
Qy	475	-----	492	-----	492
Db	31499	TCGTACGACAAATGTGTTGTTGGAGCCGATGTATCTGGGACTTTGGAGATCTTCG	31440		

Qy	492	gCyValGluAurghHisCysSerThrPheGlnMetProIlysAsp	-----ValIleil	509
Db	31439	CTGTGAGGGAACCTGTGTGCGAGGATCCGGGAAGACCGGCAGATGGTCGCCAGATTGCACG		31380
Qy	507	-----	-----ValIleil	509
Db	31379	CAGCTATGACGACGAGCTCGGAGGTGTACTTCCTGGCTGCAATCGTCTGGCTACATCTCTGAT		31320
Qy	509	eSerProHisAen	-----	513
Db	31319	CAATCCGGACCCATTACATGATACGCGAGCCAGAGTGCAGAGTGCAGAGTGCATCAAGCCTTTGGG		31260
Qy	513	-----	-----	513
Db	31259	ATTAAAGTTCCGGCAGGATTTCGGATTCGGATTCGGCCATCAATGCCACTCGGAGCGACCAATTA		31200
Qy	514	-----	-----CysgIlysgln	517
Db	31199	CGAGGCCAAGAACATCTCGTCTCMACTCGGCCACTGGCTGGTGTGGCAGCAGGAGGCCTT		31140
Qy	517	-----	-----	517
Db	31139	CACCTATGTGAGCTGGATCTGGTTCAGATCTATCGAGTCAAGGCGATTCTGTGTGAAGGG		31080
Qy	517	-----	-----	517
Db	31079	TGTGTTTACCAACGACATTGTGGCAGGCCACGGAGATTTCGTTCTTCTTACAACAAGC		31020
Qy	518	-----	-----ProAl	519
Db	31019	TGAGAGCGAGAACTACGTGGTGTACTTCTCCCAATTTCAATCTGACCATGCGAGATCCAGG		30960
Qy	519	alysPheGlyThrIleCysIyrValSer	-----CysaArgIlnGlyPheIleLe	535
Db	30959	CAACTACGGCAGCTGGCCCATGATCAGCTGCCCAAGTTCTGTGCAAGCTCGCTTGTGAT		30900
Qy	535	uSerGlyVallys	-----GluMetLeuArgCysTh	545
Db	30899	CCTTGGAAATAGTGAGCTACATGACACAGCCCTGTCTGAAGTTTCGAGTTGATGGCTGCCGA		30840
Qy	545	rThrSerGlylys	-----TrpAsnValGlyValGlnAlaValCysIy	560
Db	30839	GGAGCCGAACAGGAACCACTCTCGGCTACGACTACGGCTAC		30786
Qy	560	sAspValGluAlaProGlnIle	-----AsnCysProIlysAsp	578
Db	30785	GGACAACAGACCCATCTTCCAAATCTGCCGAGCAACCAATTTGTGTGCGACCGCA		30726
Qy	578	uGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGl		598
Db	30725	TGAGAACTGGAGGAGTACTACCGGTTAACTTACCGAACCCACCGCGGTGACAACTCCGG		30666
Qy	598	yGluIysValSerValHisValHisProAla	-----PheThrProProTyrlleuPhePr	616
Db	30665	ATCGATTCCCGCGCTGGAGATCAACGCCACAGAACTTCCGCACACCCAGCTACATTTTCAA		30606
Qy	616	oIleGlyAspValAlaIleValIleValThrAlaThrAspLeuSerGlyAsnGlnAlaSerCy		636
Db	30605	G-----GATACGGTTGAAAGTACGTGGCTTTGACTACGATGGCAATGTGCCCATCTG		30552
Qy	636	sIlePheHisIleLysValIleAspAlaGluProValIleAspTrpCysArgSerPr		656
Db	30551	CGAGATCAACATCAGGTGCCCGATGTAACACCCACTGCTGCAG		30503
Qy	656	oProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSe		676
Db	30502	CCCGACAGCTATGTGATTTGAGTTGAGCTAGTGTGATGCCGAGGACAGCTACACTGTGTGAACCTCAA		30444
Qy	676	rAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu		694
Db	30443	CGATACCCGGGAAGAGG-----ATCAAGACCTCCGACGACACAGGAGTGTGAGTTT		30393
Qy	695	-----PheProGlnGlyGluThrIle	-----ValGlnIlyrTh	705

```
Db 30392 GCAGTTACGCCCGAGAGTGCACATCAAGATCGGAACCTTGGAGAACTGACCTCTAC 30333
Qy 705 rAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleGlySe 725
Db 30332 GCGAAGGATAGATCAACAACCGCCCGCTGCCACTCCAGGTCTCTGTGAAGCTTC 30273
Qy 725 rProCys-----GluIleProPheThrProValAsnGlyAspPheIleCysThrPr 742
Db 30272 ACCCTGCGTGGAGCTCCAG---CCGCCGGGAATGGTCCCATCAATTGCTGCTGCC 30216
Qy 742 oAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGl 762
Db 30215 TGGTGTATCGTGTATCGAATGATTCACAGCTGCAAGCCAGGATTCGGTTTACCAGCG 30156
Qy 762 ySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyrThrTh 782
Db 30155 CGAACCACTGAAGACCTTCTCTCGGAGACATACGCTCTGTGGCGTCCACGCTCGTGT 30096
Qy 782 rGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLys---SerPheGl 801
Db 30095 G-----CCCGACTGCTGTGCGAGAACACCGAGCAGCGCGCTACCAGTGACCGCCTC 30042
Qy 801 uMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAl 821
Db 30041 CATTACCTACCGCGCAATGGAGCAGTGGCCCAATCTCTCTGGTTCAGTACCAGGAGT 29982
Qy 821 aPheGluThrThrLeuGly-----LysMetValProSerPheCysSerAsp----- 836
Db 29981 GCTGGCAGACCATATGGCGGACTCAACAGTTGTCTTCGACGCGTCTCGCCCGTGAA 29922
Qy 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuTh 849
Db 29921 TGTCAACATGAATGACCTTGTGTGAAGTCTGTGCCCATGCTGCTGGAGAGATGTGT 29862
Qy 849 rLysLysTyrCysLeuGluTyrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGl 869
Db 29861 CAAGATG----- 29855
Qy 869 yGlyTrpGlyAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGl 889
Db 29854 -----GACTTCATCTCTCCATT-- 29837
Qy 889 nGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLe 909
Db 29836 -----CTGCCCGCTGTGCGTTCAGCCGACGCTGTACGACCT 29802
Qy 909 uSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspGl 929
Db 29801 GTGCGGCTCCACGCTGAACCTTGATCTTTGATCTG-----AGTGATACCTATGCCAGTGC 29748
Qy 929 u---ArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeuGlnThrLeuGluTh 948
Db 29747 COTGATCGATGATGACCTTTTGAACATTTGCCAACATCGGTAAACAGTGTCTCCGCTACGCG 29688
Qy 948 rIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAl 968
Db 29687 CCTCAAGTCGGAANCTCCGAGGATTTAATCTGCAAT-----GT 29649
Qy 968 aSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysLysAlaSerProPheCy 988
Db 29648 GGGCGAGGTACTGMACTGACACACCGAGCATGTGCGC----- 29612
Qy 988 sArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTy 1008
Db 29611 -----CGTTGCTGCACTGTCCCGCCGCGAAGCATGTGT 29580
Qy 1008 rAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGluGl 1028
Db 29579 GTACAGGGTTCAGACAGCTGACCTACTGCGGAGGGGCTACTACAGAACCGTGACCG 29520
Qy 1028 yGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAs 1048
Db 1028 yGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAs 1048
Db 29519 CCAGGAACTGCTGCTGCTGCCCGCGGAACTCAACCAAGAGGAGGCGCACCAAGTC 29460
Qy 1048 nIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluTh 1068
Db 29459 GCAAGCGGACTGCATTCCTCCGCTGTATGACACCTACTACCCACCGACTGTGTGCC 29400
Qy 1068 rCysGluSerCysProLeuGlyThrTyr-----GlnProLysPheGlySerArgSe 1085
Db 29399 GTGCTTGGAGTGTCCCGTAATCTATTCTACTCGGAACCAACCAACCGTGGATTCAAG 29340
Qy 1085 rCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValaAsnIleSerAl 1105
Db 29339 TTGCCAGCTGTTCGGCACAGAGCTTCACTACCAAGCGGCTGCTCGAACAGGATCT 29280
Qy 1105 aCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisPr 1125
Db 29279 GTGTCCGCGCAAGTGTGCGCGGAAACGTACTCCGCCACCGAGACTGGCACCTCTCGCC 29220
Qy 1125 oCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPh 1145
Db 29219 CTGCCACTGCATCATATCCAGGAGCGCGGTGCGCAGAGCTGCAACGAGTGTCCGAG 29160
Qy 1145 eTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSerSe 1165
Db 29159 TAACATGAGAACCGATTACCCGCTTCAAGGAGCGGACAGTGC----- 29114
Qy 1165 rThrPheSerAlaAlaGluGluSerValValProAlaSerLeuGlyHisIleLysLy 1185
Db 29113 -----AGCCGCTGTGTA----- 29102
Qy 1185 sArgHisGluIleSerSerGlnValPheHisGluCysPheAsnProCysHisAsnSe 1205
Db 29101 -----TGTGTGAAGCTCTTCCACGACCG 29076
Qy 1205 rGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGl 1225
Db 29075 CGGACTGTGTGTGCCCATGAGCATCATCATCCAGTGTCTTGTCTCGGCGCGATTTCTCTGG 29016
Qy 1225 yLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnGlyVa 1245
Db 29015 AGCTGCTCGGAACAGACATCGACAGTGCCTTCCAGCCCTGCTTACATGTGTGTCA 28956
Qy 1245 lCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGlnAr 1265
Db 28955 GTGCAGAGATCTGCGCCAGGGCTATCGTGTAGTGTCCCGCTGGATCTCGGGCATCAA 28896
Qy 1265 gCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLysGlyIleCysVa 1285
Db 28895 TTGCCAGGAGGAGGCGCAGTGTGTGCAACGACACCTGTCTCGGCCAGGCGCATGTCAA 28836
Qy 1285 lAsp-----GlyValalagly----- 1290
Db 28835 GAACGAGCCGGCTACAGAACGTGACCTGTCTGTGCGCGAGTGTGTACACCGCGCATCA 28776
Qy 1290 ----- 1290
Db 28775 GTGCGAGTGACCATCGATCCGTGACGCGGAATGCAATCCGTGCGGAACCGAGCCAG 28716
Qy 1291 -----TyrArgCysThrCysValLysGlyPheValGlyLe 1302
Db 28715 CTGCCAGGCTTTGGAGCAGGGTCTGCTACAGTGTGTGTGCGCGCATGTGCGGAGGCGAT 28656
Qy 1302 uHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaValCy 1322
Db 28655 CCACTGTGAGCAGATATCATATGCTGTTCGGAAGATCCCTGCTGTGTGGCGCGCAACTG 28596
Qy 1322 sGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThrArgCy 1342
Db 28595 CACAGATCTGTTCAATGACTTCCAGTGTGCTGTCCGCGCAGGATTTACGGCGAGCGATG 28536
Qy 1342 sGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAs 1362
Db 28535 CGAGCAAAAGATCGATCTCTGCTATCGGAACCATGCAAGCATGCG---ACCTGCGTGA 28479
```

QY	1362	pelyAlahensSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLeu	1382	QY	1703	aaspaspPheTyrAlaGlySerThrValThrTyrGln	1715
DB	28478	TCGTCGTTCGATCAGGAGTGTGTTGGCCATCCGGCTGGACGGATCCGCTGGACAT	28419	DB	27752	CGAT	27699
QY	1382	uasnlleangluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLeu	1402	QY	1716	-CysAasnangluTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAasnGlySe	1735
DB	28418	CAACATCAGCAGCTCGGAGAACCGACCCCTGCCCAATGAGGAACTCGCTGCATCGT	28359	DB	27698	CTGTCAAGATGGC	27663
QY	1402	uasnseryrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu-ThrGluG	1422	QY	1735	rTrpAangluValSerProSerCys	1745
DB	28358	CGACGGCTATAGCTCAACTGTGAACCCCGCTACACGGCGCAAGATTCCACACACCAT	28299	DB	27662	T	27609
QY	1422	InSerThrGlyPheAasnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuA	1442	QY	1745	pValaspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAasnValas	1765
DB	28298	CGACGACTCGC	28288	DB	27663	CGCTACAGCTGCCAGTCCCTGCTTACCGGTGCAATTGCGAACAGGA	27552
QY	1442	spGlyMetLeuProSerLeuHisAlaLeuThrCys	1458	QY	1765	polySeryrileCysSerCysValProTyrThrGlyAspGlyLysAasnCysAlaG	1785
DB	28287	-----CCTCGAATCCCTGCCAGCAGCCGCCACCTGTGTGACCACTGTGATGG	28239	DB	27551	CAACGGCTTCTACTGTGCTGTC	27498
QY	1458	ysSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAasnGlySerA	1478	QY	1785	uProIleLys	1788
DB	28238	CTTTCAGTCGAATGCCCGCTGGTACTGTGGTCTCTCTCTGCG	28195	DB	27497	GGCCATCCAA	27488
QY	1478	spAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAasnGlyArgGluL	1498	RESULT 21			
DB	28194	-----AGGCCGA	28188	AC007977/c			
QY	1498	yslleThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAlaIleThrTrpT	1518	LOCUS			
DB	28187	GATCAGCAGGATCTCGAGCAGCCCTGCATCCGGTGGGACGG	28144	DEFINITION			
QY	1518	hrSerAlaasnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyValaG	1538	AC007977			
DB	28143	--AGCGCTCGCTCGATCTGGACACAAAT	28101	174287 bp			
QY	1538	lyLeuSer-ValGlyLeuProIleProGlyGly	1552	Drosophila melanogaster, chromosome 2L, region 27C-27C, BAC clone			
DB	28100	GGACGATTCAAGGACCCCTGTGGCCACGACATCGATGACTCGAGGGCGCAGCCGTG	28041	BACR1307, complete sequence.			
QY	1567	-Leu-----GlyGlnGluGlnAspLysGlyGlu	1567	AC007977			
DB	28040	TCTGAACAACGGCATCTGTTCGGATCGCTCGTGGCTTTGAGTGGCTCGACCCAGG	27981	HTG.			
QY	1567	aGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAspTyrValLeuSerProG	1587	AC007977.12			
DB	27980	A-----TGAGTGGCATCGCTCGAGCA	27957	GI:13549313			
QY	1587	nGlnVallysserLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAl	1607	Drosophila melanogaster (fruit fly)			
DB	27956	GCAGGTGACCATCGGAGCTCAGCGCGCTGCCAGAACGATGCCAGCTGCATCGACCT	27897	Drosophila melanogaster			
QY	1607	atrProAaspPheLeuSerGlyIleValGlyLysValLysIleAspSerLysSerIlePh	1627	Eukarya; Endopterygota; Insecta; Pterygota;			
DB	27896	GTTCACGACTAC	27882	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
QY	1627	eCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAs	1647	Ephydroidea; Drosophilidae; Drosophila.			
DB	27881	CTGCGTG---TGTCCAGC---GGACCATGGCAAGAACTCGAGACCGCTCCCGAA	27830	1 (bases 1 to 174287)			
QY	1647	pleulysProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAs	1667	2 (bases 1 to 174287)			
DB	27829	-----CGCTGCATCGGTGA	27816	Celniker, S.E., Aqbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,			
QY	1667	nProValGlnTyrCysLeuasnGlnGlyGlnTrpThrGlnProLeuProHisCysGlu	1686	Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,			
DB	27815	TCCT-----TGATGCAGCGGGCAAG	27796	Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,			
QY	1687	-----ArgIleSerCysGlyValProProProLeuGluAsnGlyPheHisSerAl	1703	Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,			
DB	27785	CTTTGGCTGTCTTAACTGCAGTTCCT	27753	Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,			
				Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,			
				Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,			
				Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and			
				Rubin, G.M.			
				Direct Submission			
				Submitted (06-JUL-1999) Drosophila Genome Center, Lawrence Berkeley			
				Laboratory, MS 64-121, Berkeley, CA 94720, USA			
				On Apr 5, 2001 this sequence version replaced gi:6015173.			
				Sequence submitted by:			
				Berkeley Drosophila Genome Project			
				Lawrence Berkeley National Laboratory, MS 64-121			
				Berkeley, CA 94720			
				This sequence was assembled using end sequences from a whole genome			
				shotgun and from subclones of this BAC and its neighboring clones.			

For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES

source

```

1. 174287
   Location/Qualifiers
     /organism="Drosophila melanogaster"
     /mol_type="genomic DNA"
     /strain="y; cn bw sp"
     /db_xref="taxon:7227"
     /chromosome="2L"
     /map="27C-27C"
     /clone="BACR13J07 (D857)"
     /clone_lib="BACR13J07 (D857)"
     /clone_11b="BACR13J07 (D857)"
     /notes="BACR13J07 (D857) was sequenced as a bridge.
     BACR13J07 extends from a minimal overlap with its distal
     neighbor DS01321 (D143) at bp 1 to BAC end at bp.
     142,141."

```

ORIGIN

Alignment Scores:

```

Pred. No.:      6 12e-75      Length:      174287
Score:          1328.00      Matches:      518
Percent Similarity: 32.44%      Conservative: 256
Best Local Similarity: 21.71%      Mismatches:  718
Query Match:      13.00%      Indels:      900
DB:              3           Gaps:         88

```

US-09-977-053-6 (1-1842) x AC007977 (1-174287)

```

QY      6  AlapheCysCysTrpGlyLeuAlaLeuVal-----SerglyTrpAla 19
DB      90459 GCACAAATGCTGCTGGGTAGTCTTATATTCAGATTACATAGTTCAATCAATCT 90400

QY      20 ThrPheGlnGlnMetSerProSerArgAsnPhaPheLeuPheProGluThrAla 39
DB      90399 TCATCTTGATCTTCTAGATGTGGCCACATCTCGGCCGATGTC-CGGAGCTGCAA 90341

QY      40 ProGlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySer 59
DB      90340 CCGCGA-----ATTGAGCCTTGGCCTCCAGTCGACACCGCTTTCGGTAGC 90293

QY      60 ArgValGlu-----ArgLeuGlyGlnAlaPheArgarg---ArgValArgLeuLeu 75
DB      90292 CTGGTCAGCTTTACATGTCCCATTTGGACAGAGTTTGCCACCGCGCAAGCGCACTGGTT 90233

QY      76 ArgGluLeuSerGluArgLeuGluLeuValPheLeuValAspSerSerValGly 95
DB      90232 ACCGAATG-TCTCGCGGTGGCAACTGGAGTGCTCTCTACATACCCCAAGTGTGAG--GGT 90176

QY      96 GluValAsnPhe----- 99
DB      90175 GAGTGTGCCTACCTATTCCATAATTGGATTCCATAATTCAGATATGCAATCCCTTCCA 90116

QY      100 ---ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThr 118
DB      90115 CAGAGGTCTACTCGGTCCTGTGCCAATAATGACACACCGTTTCTCCATGGCTCTCGA 90056

QY      119 AlaThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAsp 138
DB      90055 ACCTAACCTATCCGGTATAGCAATGATACAGTGCTACGCCGGCTTGGCTTCCCTCGC 89996

QY      139 TyrIleSerThrArgAla-----ArgGlnHisLys 149
DB      89995 GTGCTCCGATCGAGAGATC-TCCTGCTGCCGATGGCGCTTGGGAGCGACAGCCCAAC 89937

QY      150 Cys---AlaLeuLeuGlnGlnIleProAlaIleSerTyr----- 162
DB      89936 TGCATGGCTCTCCAGTGGCGAGCGCTGCCGGAAGTGGCACACACGCCACCGTCCGCTG 89877

QY      163 ArgGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeuLeuHis 182

```

```

DB      89876 AATGGAGTGGT----- 89865
QY      183 AlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGly 202
DB      89864 ---CGCAGCTACGCGCACCATTTGTCAGTATGAGTGTGAGCGCGCTACGAGCGCAATGCG 89808
QY      203 AspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGly 222
DB      89807 CATCCC-----GTGCTGACTGTATGTGGAACGGC 89778
QY      223 IleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHis 242
DB      89777 ACCTGGAGTGTGATGTACCAAGATGCAAG-----CGCAAGCGG 89739
QY      243 CysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAla----- 256
DB      89738 TGCTTC-----GAATTCGACCATTTGCCAACCGCTTTGTGTGGAC 89697
QY      257 ---ArgArgAlaLeuHisGluAsp-----LeuProSerGlySer 268
DB      89696 TCGACGCGAGCTTACCTCTTCGCGCATGAGGCCAGGTGTCAGTGTCTCAAGGGCTACAAA 89637
QY      269 PheIleGlnAspAspMetValHisCysSerTyr-----Leu 280
DB      89636 CTGATCGCGCAGCAACATCATCGCTCGCAGCGAGGCCAGAAAGTTTCGAGCAGCCCGCAGC 89577
QY      281 CysAspGluGlyLysAspCys-----CysAspArgMetGlySerCysLysCys 296
DB      89576 TGCAGGACATCAACGAGTGCAGCTCTCCAGTGGCGAC-----CTAACCAACCAACCGAGTGC 89520
QY      297 GlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly----- 313
DB      89519 CAGAACAGAACGGCTCTTCCACTGCAGTGCAGGACGGAGTTTCAC-GGTACCAACCGA 89461
QY      313 -----LysGlyLeuGlnTyrGluCysThrAla----- 322
DB      89460 GTGTGGCGCGCTGCTGATTTGGCTTGGTAAATGAGGCATACCGGATGACAGCATCAC 89401
QY      314 -----LysGlyLeuGlnTyrGluCysThrAla----- 322
DB      89400 CACCTCGGTGAGTGCAGCGCGGCTACAGCAAGGACAGCAGCTGCGCTTGAACACGAATGGCTG 89341
QY      323 ---CysProSerGlyThrTyrLysProGluGlySer--- 333
DB      89340 GTGCGGTGGCTCTTCGAGGCTGTGTCCAACTGGATACTCATCGACTGAAGGACACCCAC 89281
QY      333 ----- 333
DB      89280 CATTTCTGCGTGGCTTCCGCAACATGTCTCGTGCAGCGTCCCGATGCAATGTGGCTTCAG 89221
QY      333 ----- 333
DB      89220 CTCGGCGGTGCTGTGCACTACACCAACGATCTGACGAGATGTGTTCAAGGATTAATGCCAA 89161
QY      333 ----- 333
DB      89160 TCCGACGCGCATCTGCGCTGCAATTCGCAATCTGAGGCCCAAGCTCTCTCACTTTAAACCT 89101
QY      334 -----ProGly----- 335
DB      89100 GCCCTGCCCATCGNAGCTGCTATATTCGCTTCGCGATCCAGCATCTACGTGGGTGCGCC 89041
QY      336 -----Gly-IleSerSerCys-----IleProCysProAspGluAsnHisTh 349
DB      89040 CTGTCTGCGCATGAGCTGATGGCTGCAACGCGCTGTGATTTGCTGGACATCAACAGTGT 88981
QY      349 rSerProProGlySerThrSerProGluAspCysVal----- 361
DB      88980 CAGCAAGAACATGGCGGCTGTGACCAAGAGTGCATCACTCACCGGGCGGATTTGCGCTG 88921
QY      362 ---CysArgGluGlyTyrArg----- 367

```

Db 88920 TGGCTGCAACACTGGCTACCAGCTGTACACCTCCCAACGGCAGCGCTGGCTATCACAACGA 88861
Qy 368 -----AlaSerGlyGln-----ThrCysGluLe 375
Db 88860 ACGCTCCGAATCCGGCAACGCTGATGCTGACACCTATCAGCGCAACAAGACCTGTGTCC 88801
Qy 375 uValHisCysProAlaLeuAspProGluAsnGlyTyrPheIleGlnAsnThrCysAs 395
Db 88800 TCTCATGTGTCCGAACCTGGAGCGCCCGAGAAATGGTCAACTCTCAGCGACACAAGACGA 88741
Qy 395 nAsnHisPheAsnAlaAaCysGlyValArgCysHisProGlyPheAspLeuValGlySe 415
Db 88740 CTATCATTTGGGATGTGGTGGCTTCAGTGCCACTTTGGCTACATCATGAGCGGAG 88681
Qy 415 rSerIleLeuLeuCysLeuProAsnGlyLeuTrpSerGlySer-----GluSerTyrCy 433
Db 88680 CTGGCGGCGCTCTCCAGCGGTCAAGTGAACCGCAGCGTACCGGAGTGCATTTG 88621
Qy 433 sArgValArgThr----- 437
Db 88620 TAAGTGTATATCATTTTCAATTCATGAATTTCCATTTTACTAAACGAATAATGTCAC 88561
Qy 437 ----- 437
Db 88560 TTTTCAGATGCCAATTCGTTTCCCTGCCGATGACAAGTTGAGGGTCTGACTGTGGCC 88501
Qy 438 -----CysProHisLeuArgGln----- 443
Db 88500 CGCCCGATCCCGAATCCGTTCTAGTGCCTTCGCGTGAACAATGTGACCATTTAGCTGCGGA 88441
Qy 443 ----- 443
Db 88440 TCGCCGGGAGCCAACTGAGAGCCACCGCTTCTCTGGTTTCGGCAGTGCCTGTACGAT 88381
Qy 444 -----ProLysHisGlyHisLeuSerCys----- 451
Db 88380 CCCAAGCCCGTCTGCGCAATTAAGTCTGCTATCGCGATGAGCCCTCTTGTCCCGAGTG 88321
Qy 452 -----SerThrArgGluMetLe 457
Db 88320 GATTGCTACTACCCATGCCAAGCCCGCGCAGATACGACAGTGTGTG-CACACTCG 88262
Qy 457 uTyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgGluGly----- 474
Db 88261 CTATCAGACGACTTCTTTTGGTGCAGAACACTTTTAAGTTGGTGGCAGACAGCGG 88202
Qy 475 -----SerAspLysLeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProAr 492
Db 88201 TCGTCACGACATGTGGTGTGGAGCCGATGTATCTGGACTTTGGAGATCTTCG 88142
Qy 492 gCysValGluArgHisCysSerThrPheGlnMetProLysAsp----- 506
Db 88141 CTGTGAGGACCTGTGTGAGGATCGGAGATCGGAGACCGGAGATGGTGGCCAGATGTCACG 88082
Qy 507 -----ValIleI 509
Db 88081 CAGCTATGACGAGCTCGGAGGTGTATCTTCGGCTGCATTCGCTCGGCTATCATCTGAT 88022
Qy 509 eSerProHisAsn----- 513
Db 88021 CAATCCGACCCATATACATGATACGCGGACGAGTGCAGAGTTCATCAAGCTTTGGG 87962
Qy 513 ----- 513
Db 87961 ATTAAGTTCGGCAGGATTCGGATTCGGCCATCAATGCCACCTCGAGCGACCCAAATTA 87902
Qy 514 -----CysGlyLysGln----- 517
Db 87901 CGAGGCCAAGAACATCCGTTCTCAATCGGCCACTGGCTGTGTGGCAAGCAGAGGCCTTT 87842
Qy 517 ----- 517
Db 87841 CACCTATGTGAGCGGTGATCTGGGTTCAGATCTATCGAGTCAAGGGCGATTTCTGGTGAAGGG 87782

Qy 517 ----- 517
Db 87781 TGTGTTTACCAACGACATTTGTGGCAGGCCCAACGAGATTCGGTTCTTCTACAACAAGC 87722
Qy 518 -----ProAl 519
Db 87721 TGAGAGCGAGAACTAGCTGTGTACTTCTCCCAATTTCAATCTGACCATGCGAGATCCAGG 87662
Qy 519 aLysPheGlyThrIleCysTyrValSer-----CysArgGlnGlyPheIleLe 535
Db 87661 CAACTACGGCGAGCTGGCCATGATCAGCTGCCCAAGTTCTGTGAGGCTCGCTTTGTGAT 87602
Qy 535 uSerGlyValLys-----GluMetLeuArgCysTh 545
Db 87601 CCTTGGATATAGTAGCTACATGACCAACGCTGTCTGTAAGTTTCGAGTTGATGGCTCGCA 87542
Qy 545 rThrSerGlyLys-----TrpAsnValGlyValGlnAlaAlaValCysLy 560
Db 87541 GGAGCGCAACAGGAACCACTCTCTGGCTAGCACTACGCTAC-----TCCCGTGGGT 87488
Qy 560 sAspValGluAlaProGlnIle-----AsnCysProLysAsp-----IleGluAlaLysThrLe 578
Db 87487 GGACCAACGACCACTCTTCCAAACTGCCCGCAGCAACCAATTTGTGTGGCAGCGCA 87428
Qy 578 uGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGl 598
Db 87427 TGAGAATGAGGAGTACTACCGCTTAACCTTCAACCGAACCCCGCGGTGGCAACTCCGG 87368
Qy 598 yGluLysValSerValHisValHisProAla-----PheThrProProTyrLeuPhePr 616
Db 87367 ATCGATTGCCCTGGAGATCAAGCCACAGAACTTCCGACACCCAGCTACATTTTCAA 87308
Qy 616 oIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCy 636
Db 87307 G-----GATACGGTTGTAAGTACGTGGCTTTGACTACGATGGCAATGTGGCCACTG 87254
Qy 636 sIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSerPr 656
Db 87253 CGAGATCAACATCAGCGTCCCGATGTAAACACCACCACTGCTGCAG-----TGC----- 87205
Qy 656 oProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSe 676
Db 87204 -CCCCAGACTATGTGATGTAGCTAGTGGATCGCCAGGACGACTACACTGTGAACTTCAA 87146
Qy 676 rAspAsnSerGlyAlaGluValIleThrArgSerHisThrGlnGlyAspLeu----- 694
Db 87145 CGATACCCGGAAGAG-----ATCAAGACTCTCCGACGACACAGGAGATGTAGGTT 87095
Qy 695 -----PheProGlnGlyGluThrIle-----ValGlnTyrTh 705
Db 87094 GCAGTTTACGCCCGAGAGTGCCCAACATCAAGATCGGAAACTTCGAGAACGTGACCGTCAC 87035
Qy 705 rAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySe 725
Db 87034 GGCAACGATATAGTACACAAACCGCGCCCTGCCACTTCCAGGTCTCTGTGAGAGGCTTC 86975
Qy 725 rProCys-----GluIleProPheThrProValAsnGlyAspPheIleCysThrPr 742
Db 86974 ACCCTGGTGGACTGGAGCTCCAG-----CCGCGCGCAATGGTGCCTCAATTTGCCCTGCC 86918
Qy 742 oAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThrGluGl 762
Db 86917 TGGTATCTGGTATCGAATGCAATTCCTCCAGTGCAGCGCAAGCCAGGATTCGTTTACCGACGG 86858
Qy 762 ySerThrAspLysTyrCysAlaTyrGluAspGlyValTrpLysProThrTyrThrTh 782
Db 86857 CGAACCACTGAAGACCTTCTCTCGGAGACATCAGCTCTGTGGCTGCCAGCTCGGTGT 86798
Qy 782 rGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLys-----SerPheGl 801
Db 86797 G-----CCGCACTGGTGTGCGAGAACACGAGGAGCGCGCCTTACCACGTGACGCTC 86744

QY 801 uMetPheTyrLysAlaalaArgCysAspThrAspLeuMetLysLysPheSerGluAl 821
Db 86743 CATTAACCGCGCCCAATGAGCAGTGGCCCAATCTCTGTGGTCTGACAGGAGGT 86684
QY 821 aPheGluThrThrLeuGly-----LysMetValProSerPheCysSerAsp----- 836
Db 86683 GCTGGCACACACTATGGCGGACTCAACACAGTGTCTCTCGCAGCGCTGCTCCGCGGTAA 86624
QY 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuTh 849
Db 86623 TGTCAACATGAATGTGACCTTTGTGAAGTCTGTGCCCATCTGCTGGAGAGAAATGTGT 86564
QY 849 rLysLysTyrCysLeuGluThrAsnTyrAspTyrGluAsnGlyPheAlaIleGlyProGI 869
Db 86563 CAAGATG----- 86557
QY 869 yGlyTrpAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGI 889
Db 86556 -----GACTTCATCTCTCCATT-- 86539
QY 889 nGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLe 909
Db 86538 -----CTGCCCGCTGTGCGTCAACATCGGTAACTGCTCTCCGCTACGCGC 86504
QY 909 uSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspGI 929
Db 86503 GTCCGCTCACCGCTGAACCTGATCTTTGATCTG-----AGTGATCCCTATGCCAGTGC 86450
QY 929 u--ArgAsnAspThrLeuGluThrGluAsnGlnGlnArgLeuLeuGlnThrLeuGluTh 948
Db 86449 CGTGATCGATGACCTTTTGAACATTCGCAACATCGGTAACTGCTCTCCGCTACGCGC 86390
QY 948 rIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTyrSerPheGlnLeuAl 968
Db 86389 CCTCAAGTGCCTCAATCTCGCGGATTTAACTGCAAT-----GT 86351
QY 968 aSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaAsrProPheCy 988
Db 86350 GGGCGAGGTACTGAACATGACACACCGATGCGCG----- 86314
QY 988 sArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTy 1008
Db 86313 -----CGTGGCTGCACTGTCTCGCGCGGAGTATGT 86282
QY 1008 rAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrGlnAspGluGluGI 1028
Db 86281 GTCCAGAGGGTCAAGACAGCTGCACCTACTCCCGAGGGGCTACTTACCAGAACCGTGCACCG 86222
QY 1028 yGlnLeuGluCysLysLeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAs 1048
Db 86221 CCAGGGAACCTGCTGGCTGCGCGCGCGGACCTACCCAGAGAGAGGGGCGACCAAGTC 86162
QY 1048 nIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluTh 1068
Db 86161 GCAGCGGAGTCAATCTCCGCTGCTGGGTATGCACTACTCACCCACCGAGTGGTGC 86102
QY 1068 rCysGluSerCysProLeuGlyThrTyr-----GlnProLysPheGlySerArgSe 1085
Db 86101 GTCCCTGGAGTGTCCGGTAACCTCATCTACGCGCAACCAACCGGTGGATTCAAGGA 86042
QY 1085 rCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAl 1105
Db 86041 TTGCCAGGCTGTCCGGCACAAGAGTCTACCTACCGCGCGCTGCTTCGAAACAAGATCT 85982
QY 1105 aCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisPr 1125
Db 85981 GTGTCCGGCAAGTGTGGCGCGGAAAGTACTCCGCCACCGAGTGCACCTGCTCGCC 85922
QY 1125 oCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPh 1145
Db 85921 CTGCCCACTGCATCATTAACCAAGAGCGCGGGTGGCGAGAGTGCACAGTGTCCGAG 85862
QY 1145 eTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSerSerPheSerSe 1165

Db 85861 TAACATGAGAACCGATTTCACCGCTCCAGGGAGCGGACAGTGC----- 85816
QY 1165 rThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGlyHisIleLysLy 1185
Db 85815 -----AAGCGGTGGTA----- 85804
QY 1185 sArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsnProCysHisAsnSe 1205
Db 85803 -----TGGTGAAGGTGCTTCCACACGCG 85778
QY 1205 rGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysProLeuGlyTyrThrGI 1225
Db 85777 CGGACTGTGTGTCCTCCATGGCCATGACATTCAGTGTCTTCTCGCGCGGATTTCTCTGG 85718
QY 1225 yLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGlyVa 1245
Db 85717 ACGTGCGTGGCAACAGGACATCGACAGTGTGCGCTCCAGCCCTCTACAAATGGTGTCA 85658
QY 1245 lCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrThrGlyGluAr 1265
Db 85657 GTGCAAGATCTGCGCAGGCTATCGTGTGAGTGGCCGCTGGATCTCGGGATCAA 85598
QY 1265 gCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsnLysGlyIleCysVa 1285
Db 85597 TTGCCAGGAGGCGCAGTGTGCTGGCAACGACACCTGTCCGCCAGGGCCATGTGCAA 85538
QY 1285 lAsp-----GlyValAlaGly----- 1290
Db 85537 GAACGAGCGGGCTTACAAGAACGTGACCTGTCTGTGCGCAGTGTGCTACACCGCGATCA 85478
QY 1290 ----- 1290
Db 85477 GTGCGAGTGCACATCGATCGTGCACGCGGATGGCAATCCGTGCGGAAACGGAGCCAG 85418
QY 1291 -----TyrArgCysThrCysValLysGlyPheValGlyLe 1302
Db 85417 CTGCCAGGCTTGGAGCAGGCTCGCTACAAGTGTGAGTGTGTGCGCGATGGAGGCGAT 85358
QY 1302 uHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCy 1322
Db 85357 CCATGTGTGACGAGAATATCAATGATCTGTTCGAGAAATCCCTGCTGTTGGGCGCCAACTG 85298
QY 1322 sGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCy 1342
Db 85297 CACAGATCTGCTCAATGACTTCCAGTGTGCTGCGCCAGATTTACGGGCAAGCGATG 85238
QY 1342 sGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLysAs 1362
Db 85237 CGAGCAAAAGATCGATCTCTGCTATCGGAACCATGCAAGCATGCG--ACCTGCGTGA 85181
QY 1362 pGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLe 1382
Db 85180 TCGTCTGTTCGATCAGCAGTGTGTTTTGCCATTCGCGGCTGGAGCGGATCCGCTCGACAT 85121
QY 1382 uAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLe 1402
Db 85120 CAACATCGAGACTGCGAGAACCGACCTGCGCCATGAGGGAACCTCGCTCGACTCGT 85061
QY 1402 uAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu-ThrGIUG 1422
Db 85060 CGACGGCTATAGTGCACACTGTGAACCGGCTACACGGGCAAGAAATGGCCAGCACCAT 85001
QY 1422 lnsThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeuA 1442
Db 85000 CGACGACTGCG----- 84990
QY 1442 spGlyMetLeuProSerLeuHisAlaLeuThrCys-----ThrPheTrpMetL 1458
Db 84989 -----CCTCGAATCCCTGCGCAGCAGCGGCCACCTGTGTGGACCGAGTGGATGG 84941
QY 1458 ysSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerA 1478

84940	Db	CTTCAGCTGC	AAATGGCGCCCTGGCTACGCTGGGTCTCTCTCTGGC	----	84899
1478	Qy	spAsnThrLeu	LeuLeuThrAspTyrAsnGlyTyrValLeuTyrVal	AsnGlyArgGluL	1498
84896	Db	-----	-----	-----AGGCCGA	84890
1498	Qy	ysileThrAsn	CysProSerValAsnAspGlyArgTrpHisIleAla	IleThrTrpT	1518
84889	Db	GATCGACAGT	GTCTCGACGACCCCTCGCAATCCGGTGGGCACGG	-----	84846
1518	Qy	hrSerAlaAsn	GlyIleTyrValTyrIleAspGlyIysLeuSerAspGlyGlyAla	G	1538
84845	Db	--AGCGTC	GCCTCGATCTGCACACAAAT	-----TCGAGTCGGTGTGCCG	84803
1538	Qy	lyLeuSer	ValGlyLeuProIleProGlyGly	-----GlyAlaLeuVal	1552
84802	Db	GGACGGATT	CAAGGGACCCCTGTGCGCCACGACATCGATGACTCGAGGCGCAGCCGTG	84743	
1553	Qy	-Leu	-----GlyGlnGlnAspIysIysGlyGlu	-----GlyPheSerProAl	1567
84742	Db	TCTCAACAC	AGGCATCTGTCCGGATCGCGTCGGTGGCTTTGAGTCGGCGCTCGCAGCCAGG	84683	
1567	Qy	aGluSerPhe	ValGlySerIleSerGlnLeuAsnLeuTyrAspTyrValLeuSerProGl	1587	
84682	Db	A	-----TCGAGTCGCATCGCGCTGCAGCA	84659	
1587	Qy	nGlnVal	SerLeuAlaThrSerCysProGluLeuLeuSerIysGlyAsnValLeuAl	1607	
84658	Db	GCAGGTGAC	ACATCGCGGAGCTCAGCGCGCTGCCAGACGATGCCAGCTGCATCGACCT	84599	
1607	Qy	atPrProAsp	PheLeuSerGlyIleValGlyIysValIysIleAspSerIysSerIleP	1627	
84598	Db	GTTC	CAGGACTAC	-----TT	84584
1627	Qy	eCysSerAsp	CysProArgLeuGlyGlySerValProHisLeuA	ThrAlaSerGluAs	1647
84583	Db	CTGCGTG	-----TGTC	CCAGC-----	84532
1647	Qy	pLeuLysPro	GlySerIysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAs	1667	
84531	Db	-----	-----CGCTGCATCGGTGA	84518	
1667	Qy	nProValGln	TyrCysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGlu	--	1686
84517	Db	TCCT	-----TGCATGCACCGTGGCAAG	-----TGCCACGA	84488
1687	Qy	-----	-----ArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAl	1703	
84487	Db	CTTTGGCT	CTGGTCTTAACTCGATGGCCCT	-----GC	84455
1703	Qy	aAspAspPhe	TyrAlaGlySerThrValThrTyrGln	-----	1715
84454	Db	GGAT	-----TACTCGGGCATTTGGTGTCTAGTACGAGTACGACGCATCGCAGGACATGT	84401	
1716	Qy	-CysAsnAsn	GlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlyse	1735	
84400	Db	CTGTGC	ACATATGGC	-----GCCACTTGTGTGCACATGTGTGC	84365
1735	Qy	rTrpAsnGly	ValSerProSerCys	-----LeuAs	1745
84364	Db	T	-----GGCTACAGCTGCCAGTGGCCCACTGGCTTCACCGGTTCGCATTTGGACACACGA	84311	
1745	Qy	pValAspGlu	CysAlaValGlySerAspCysSerGluHisAlaLeuSerCysLeuAsnValAs	1765	
84310	Db	CATCGTGA	CTGCACAGACAACTCT	-----TGCCACACCGGGCGCCACGTGGTGGTGTCTAAC	84254
1765	Qy	pGlySerTyr	IleCysSerCysValProProTyrThrGlyAspGlyIysAsnCysAlaG	1785	
84253	Db	CAACGGCT	TCTTACTGTCTCAGTGC	-----CCCTTCATATGACCGGAGACGATTCGCGCA	84200
1785	Qy	uproIle	Iys	1788	
84199	Db	GGCCAT	CTCAA	84190	

RESULT 22
AE00361579
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AUTHORS

Misra, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., Campbell, K., Hradecky, P., Huang, Y., Kaminker, J.S., Prochownik, S.E., Smith, C.D., Tupy, J.L., Bergman, C.M., Berman, B.P., Carlson, J.W., de Ceiniker, S., Clamp, M.E., Drysdale, R.A., Emert, D., Frise, E., de Grey, A.D.N.J., Harris, N.B., Kronmiller, B., Marshall, B., Millburn, G.H., Richter, J., Russo, S., Searle, S.M.J., Smith, E., Shu, S., Smutnak, P., Whitfield, E.J., Ashburner, M., Gelbart, W.M., Rubin, G.M., Mungall, C.J. and Lewis, S.E.

TITLE

Annotation of Drosophila melanogaster genome

JOURNAL

Unpublished

REFERENCE

4 (bases 1 to 270766)

AUTHORS

Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.

TITLE

Direct Submission

JOURNAL

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

REFERENCE

5 (bases 1 to 270766)

AUTHORS

PlyBase

CONSTRM

Direct Submission

TITLE

Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

JOURNAL

6 (bases 1 to 270766)

REFERENCE

PlyBase

AUTHORS

Direct Submission

CONSTRM

Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA

JOURNAL

On Sep 16, 2002 this sequence version replaced gi:2797167.

COMMENT

Location/Qualifiers

FEATURES

1. 270766

source

/organism="Drosophila melanogaster"

gene

/mol_type="genomic DNA"

mRNA

/db_xref="taxon:7227"

CDS

/chromosome="2L"

gene

/note="genotype: y[1]; cn[1] bw[1] sp[1]; Rh6[1]"

mRNA

complement(1220..2984)

CDS

/locus_tag="CG11322"

gene

/note="last curated on Thu Aug 01 11:12:30 PDT 2002"

mRNA

/map="27A2-27A2"

CDS

/db_xref="FLYBASE:FBgn0031856"

gene

complement(1220..2984)

mRNA

/locus_tag="CG11322"

CDS

/db_xref="FLYBASE:FBgn0031856"

gene

complement(1408..2886)

mRNA

/locus_tag="CG11322"

CDS

/note="CG11322 gene product"

gene

/codon_start=1

mRNA

/product="CG11322-PA"

CDS

/protein_id="AAF52434.2"

gene

/db_xref="GI:22945804"

mRNA

/db_xref="FLYBASE:FBgn0031856"

CDS

/translation="NDSRSNDTMEIQKLPVKEHLVTVKGHCQVKNCRNLDLM

gene

STQPSDSSQSEGBSDNDDEYQOPEWGRKRPPLVLEIGDKIIVRNPQ

mRNA

PVPSADDVQLKTKHFKGLIPSRKLYNPAYVDMKPEIDENVCVNSHAGAB

CDS

FNFKSQPRPKMTSTWKLNDIETPEESQAKNDPELMSAFSDQCTPD

gene

LHKPEIDLNEQLAACRRHSVCLSRHPKFGFVRLSPSISLNWCHPLAL

mRNA

TYRQNNKTCVKALAKVNIENHAFHGLQVPIVWKHAMTRCTCOLTVDAHGRR

CDS

AKILLRSIKTTAEILRCLLHEMCHVAFFVRELGHGDCRRWAYQAKALPEPTI

gene

EDNDSFPTCTMTCRGVADLSLMQRLCYIQFVSVKRWKSEIMYDWRSDPT

mRNA

MTFFKCFIWDRLYQGEQANTSHNMALLBEFRLNSTMD"

CDS

complement(3883..14000)

gene

/locus_tag="CG11321"

mRNA

/note="last curated on Thu Aug 01 11:11:34 PDT 2002"

CDS

/map="27A2-27B1"

gene

/db_xref="FLYBASE:FBgn0031857"

mRNA

complement(Join(3883..4084,4154..5049,5125..5391,

CDS

5463..7465,8264..9238,9327..9882,9942..10369,10614..11801,

gene

11860..12148,12202..12324,12754..13092,13309..13437,

mRNA

13490..13626,13691..13889,13947..14000))

CDS

/locus_tag="CG11321"

gene

/product="CG11321-PA"

mRNA

/db_xref="FLYBASE:FBgn0031857"

mRNA

complement(Join(3883..4084,4154..5049,5125..5391,
5463..9238,9327..9882,9942..10369,10614..11801,
11860..12148,12202..12483,12754..13092,13309..13437,
13490..13626,13691..13889,13947..14000))
/locus_tag="CG11321"

CDS

/product="CG11321-PA"
/db_xref="FLYBASE:FBgn0031857"
complement(Join(3883..4084,4154..5049,5125..5391,
5463..7465,8264..9238,9327..9882,9942..10369,10614..11801,
11860..12148,12202..12324,12754..13092,13309..13437,
13490..13626,13691..13889,13947..14000))
/locus_tag="CG11321"
/note="CG11321 gene product from transcript CG11321-RA"
/codon_start=1
/product="CG11321-PA"
/protein_id="AAF52435.2"
/db_xref="GI:22945805"

Alignment Scores:

Pred. No.:	1,12e-74	Length:	270766
Score:	1328.00	Matches:	518
Percent Similarity:	32.44%	Conservative:	256
Best Local Similarity:	21.71%	Mismatches:	718
Query Match:	13.00%	Indels:	900
DB:	3	Gaps:	88

US-09-977-053-6 (1-1842) x AE003615 (1-270766)

Qy	6	AlaPheCysCysTrpGlyLeuAlaLeuVal	-----SerGlyTrpAla 19
Db	226968	GCACATGCTGCTGGTAGTGTCTTATAATTAGATTACATAGTTCATCACTAATCT	-----
Qy	20	ThrPheGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAla 39	-----

DB	226908	TCAATCTTGAATCTTTAGATGGTGGGCACATTTCTCGGCCGATATGC-CCGGAGCTGCNA	226950	DB	225969	GTGTGGCCCGCTCGGTGATTTGGGCTTGGTAATGAGGACATACCGGATGACAGCATCAC	225910
QY	40	ProGlyValProGlySerProAlaProAlaProGlyAspGluAlaAlaGlySer	59	QY	314	-----LysGlyLeuGlnTyrGluCysThrAla-	322
DB	226849	CCCGA-----ATTGAGCGCTTGGCTCCAGTCGGACACCGCTTTCGTAGC	226802	DB	225909	CACCTCGGTCACTGAGCGGGCTACAGCAAGGAGCAGCTCGCTTGAACACGAATGGCTG	225850
QY	60	ArgValGlu-----ArgLeuGlyGlnAlaPheArgArg-----ArgValArgLeuLeu	75	QY	323	-----CysProSerGlyThrTyrLysProGluGlySer---	333
DB	226801	CTGGTCAGCTTACATGTCCTTGGCAGGAGTTTGGCACCGGCAAGCGACTGGTT	226742	DB	225849	GTGGGTGGCTCTTCGGAGCGCTGGTGGCACTGGATACTATCGACCTGAGGACCCAC	225790
QY	76	ArgGluLeuSerGluArgLeuGluLeuValPheLeuValAspSerSerValGly	95	QY	333	-----	333
DB	226741	ACCGAATG-TCTGCGGTGGCACTGGAGTGTCTCTACATACCACAAAGTGTACG--GGT	226685	DB	225789	CAATCTGGGTGGCTTCGGACCATGTCTCGTCAGCGGTCCCGATGGCAATGTGGCTTTCAG	225730
QY	96	GluValAsnPhe-----	99	QY	333	-----	333
DB	226684	GAGTGTGCTACTATTCCATAATTGGATTCCATAACTAATCGATATGCAATCCCTTCCA	226625	DB	225729	CTCGCGGTGCTCTGCAGTACACCAACGATCTGACGGATGTGTTTCAAGGATTTATGCCAA	225670
QY	100	---ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThr	118	QY	333	-----	333
DB	226624	CAGAGTTCTACTGCGGTCTGTGTGCCCAAAATCGACAACGGTTTCTCCATTGGCTCTCGA	226565	DB	225669	TCCGACGGCACTGCGCTCGGAATTCGGGCHATTCCTGGAGCCACGCTCTCCATCTTAAACCT	225610
QY	119	AlaThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAsp	138	QY	334	-----ProGly-----	335
DB	226564	ACGTAACTATCGCGGTATAGCAATGTACAGTCTACGCCGGCTTTCCTTCGCTCGG	226505	DB	225609	GCCCTGCCCCATCGAAGCTCGCTATATTGCTTCCGCATCCAGGACTACGTGGGTGGCC	225550
QY	139	TyrIleSerThrArgAla-----	149	QY	336	-----Gly-IleSerSerCys-----IleProCysProAspGluAsnHis	349
DB	226504	GTGCTCCGATCGAAGATC-TCTGTCTGCGGATGGCGGTGGGAGCGACAGCCAC	226446	DB	225549	CTGTCTGGCGATGAGCTGATGGGCTGCAACGGCTTGGATTGGTGGACATCAACAGTG	225490
QY	150	Cys---AlaLeuLeuGlnGluLeuProAlaIleSerTyr-----	162	QY	349	rSerProProGlySerThrSerProGluAspCysVal-----	361
DB	226445	TGCATGGCTTCCAGTCGCGCAGCTGCGGAGTGGCACACGCCAACGTCACCTCGCTG	226386	DB	225489	CAGCAAGAACAATGGGGCTGTGACCGAGTGCATCACTACCCGGGCGGATTTGCCTG	225430
QY	163	ArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaGlnIleLeuLeuHis	182	QY	362	-----CysArgGluGlyTyrArg-----	367
DB	226385	AATGGAGTGGT-----	226374	DB	225429	TGGCTGCAACACTGGCTACAGCTGTACACCTCCAAACGGCACGGCTGCTATCACATCGA	225370
QY	183	AlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGly	202	QY	368	-----AlaSerGlyGln-----ThrCysGluLeu	375
DB	226373	---CCGAGTACGGCACCATCTTCTCCAGTATGATGTGAGCGCGGTACGAGCGCAATGGC	226317	DB	225369	ACGCTCCGAATCCGGCAACGATGATGTGACACCTATACAGCGCAACAAGACCTGTGTTC	225310
QY	203	AspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGly	222	QY	375	uValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAs	395
DB	226316	CATCCC-----GTGCTGACCTGTATGTCGAACGGC	226287	DB	225309	TCTCATGTCTCCGAACTGGAGCGCCCGAGATGTCTCACTCTCTGAGCGCAAGAAGA	225250
QY	223	IleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHis	242	QY	395	nAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySe	415
DB	226286	ACCTGGAGTGGTGTATACCAAGATGCACG-----CGCAAGCGG	226248	DB	225249	CTATCACTTTGGCGATGTGTGGCTTCCAGTGCACCTTTGGCTATCATCATGAGCGGCG	225190
QY	243	CysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAla-----	256	QY	415	rSerIleIleLeuCysLeuProAsnGlyLeuTrpSerGlySer-----GluSerTyrC	433
DB	226247	TGCTTC-----GAAATCCGACCATTTGCCAAGCTTTGTGTGGAC	226206	DB	225189	CTCGCGGCGCTGTGTGCTCTCCAGCGGTGAGTGAACGCCAGGTACCGGATGCAATTG	225130
QY	257	-----ArgArgAlaLeuHisGluAsp-----	268	QY	433	sArgValArgThr-----	437
DB	226205	TGACGCGAGCTTACTCTCTCGCGATGAGCGCGGTGAGTGTCTCAAGGGCTACAAA	226146	DB	225129	TAAGTGTATATATTCAATTCAGATTTCTCAATGAAATTCCTTCAACAACGAATGTGCAC	225070
QY	269	PheIleGlnAspAspMetValHisCysSerTyr-----	280	QY	437	-----	437
DB	226145	CTGATCGGAGCAACATCATGCGCTGACGAGGCGCCAGAAAGTTGAGCAGCGCGGAGC	226086	DB	225069	TTTTCAGATGCCAAATGCGTTTCCCTGCGCGATGACAAAGTTGGAGGGTCTGACTGTGGCC	225010
QY	281	CysAspGluGlyLysAspCys-----CysAspArgMetGlySerCysLysCys	296	QY	438	-----CysProHisLeuArgGln-----	443
DB	226085	TGCGAGGACATCAACGAGTGCAGCTCTCGCAGTGGCAG-----CTAACCAACCGAGTGC	226029	DB	225009	CGCCCGATCCGAAATCCGTTCTAGTGCCTTCCGTTGACAAATGTGACCATTACGTGGGA	224950
QY	297	GlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly-----	313	QY	443	-----	443
DB	226028	CAGAACACGAACGGCTCTCTTCCACTGCCAGTGCAGACCGGATTCAC--GGTACCAACGA	225970	QY	444	-----ProLysHisGlyHisIleSerCys-----	451
DB	313	-----	313	DB	224889	CCCAAGCCGGTGTGCGCGATTTACTGGCTATCGGAATGCAGGCCCTCTTGTCCCGGAGTG	224830

```
QY 452 -----SerThrArgGluMetLe 457
Db 224829 GATTGTACTACCCATGCCAAGCCCGCGGCGAGAAATACGACAGTTTGTG-GACACTCG 224771
QY 457 uTyTyLysThrThrCysLeuValAlaCysAspGluGlyTyArgLeuGluGly-----474
Db 224770 CTATCAGACGAGCTTCTTCTTGGCTGCCAGAACACCTTTAAAGTTGGCTGGCAGACAGCGG 224711
QY 475 -----SerAspLysLeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProAr 492
Db 224710 TCCTCAGCAGCAATGTGGTTGTGGTGGAGCGGATGATCTCGGACTTTGGAGATCTTCG 224651
QY 492 CyLeuValGluArgHisCysSerThrPheGlnMetProLysAsp-----506
Db 224650 CTGTGAGGAGACCTGTGTGGCAGGATCGGGAAGACCGGAGATGGTCGCCAGATTGCAAG 224591
QY 507 -----ValIleIle 509
Db 224590 CAGCTATGAGCAGAGCTGGGAGGTGTACTTTCGGCTGCAATCGTCTCGGCTACATCTCTGAT 224531
QY 509 eSerProHisAsn-----513
Db 224530 CAATCCGCGACCCATTACATATCATACGAGCCAGAGTGCAGGTGCATCAAGCTTTGGG 224471
QY 513 -----513
Db 224470 ATTAAAGTTCCGGCAGGATTCGGGATTCGGGCATCATATGCCACCTCGGAGCGACCCCAATTA 224411
QY 514 -----CysGlyLysGln-----517
Db 224410 CGAGGCCAAGAACATCGTCTCAACTCGGCCACTCGGCTGTGTGGTGGCAAGCAGGAGGCGCTT 224351
QY 517 -----517
Db 224350 CACCTATGTAGCGTGTGATCTGGTTCAGATCTATPCGAGTCAAGCGGATTCGTGTGAAGGG 224291
QY 517 -----517
Db 224290 TGTGGTTACCAACGACATTTGTGGCAGGCCCGCCAGAGATTCGGTTCTTCTACAAACAAGC 224231
QY 518 -----ProAl 519
Db 224230 TGAGAGCGAGAACTAGCTGGTGTACTTCCGCCAATTTCAATCTGACCATCGAGATCCAGG 224171
QY 519 aLysPheGlyThrIleCysTyValSer-----CysArgGlnGlyPheIleLe 535
Db 224170 CAACTAGCGCGAGCTGGCCATGATCAGCTGCCCAAGTTCTGTCAGGCTCGCTTTGTGAT 224111
QY 535 uSerGlyValLys-----GluMetLeuArgCysTh 545
Db 224110 CCTTGGATAGTACGCTACATGACAAACGCGCTGTCTGAAGTTTCGATTTGATGGCTGCGA 224051
QY 545 rThrSerGlyLys-----TrpAsnValGlyValGlnAlaAlaValCysLys 560
Db 224050 GGAGCGGAAACAGAACCACTCTCGGCTACGACTACGGCTAC-----TCCCGTGGCT 223997
QY 560 sAspValGluAlaProGlnIle---AsnCysProLysAsp---IleGluAlaLysThrLe 578
Db 223996 GGCAACAGCAACCACTCTTCCAAACACTCCCGCAGCAACCAATTTGTGTGGCAGCGCA 223937
QY 578 uGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerG1 598
Db 223936 TGGAATGGGAGTACTACCGTTAACTTTCACCGAACCCACCGCGGTGGACAACTCCG 223877
QY 598 yGluLysValSerValHisValHisProAla-----PheThrProProTyLeuPhePr 616
Db 223876 ATCGATTGCCCGCTGGAGATCAAGCCACAGAACTCCCGCACACCCAGCTACATTTTCAA 223817
QY 616 olleGlyAspValAlaIleValTyThrAlaThrAspLeuSerGlyAsnGlnAlaSerCy 636
Db 223816 G-----GATACGTTGTAAAGTACGTGGCTTTGTACGTAGTGGCAATGTGGCCATCTG 223763
```

```
QY 636 silePheHisIleLysValIleAspAlaGluProValIleAspTrpCysArgSerPr 656
Db 223762 CGAGATCAACATCAGCGTGGCGGATGTAACACCACTGCTGCAG---TGC-----223714
QY 656 oProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSe 676
Db 223713 -CCCGAGAGCTATGTGATTGAGCTAGTGATCGCGAGCAGCTACACTGTGAACCTTCAA 223655
QY 676 xAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu----694
Db 223654 CGATACCCCGGAAGAGG-----ATCAAGACCTCCGACACACAGGAGATGTGAGTT 223604
QY 695 -----PheProGlnGlyGluThrIle-----ValGlnTyTh 705
Db 223603 GCAGTTCCAGCCCGAGAGTGCCACATCAAGATCGAAACTTCGAGAACTTCGAGAACTGCAC 223544
QY 705 rAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySe 725
Db 223543 GGCAACGGATAAGTACAAACACCGCGCTGCCACTTCAGAGTCTCTGTGAAGGCTTC 223484
QY 725 rProCys-----GluIleProPheThrProValAsnGlyAspPheIleCysThrPr 742
Db 223483 ACCCTGCTGAGCTGGAGCTCCAG-----CCGCCGCGAATGGTCCCATCAATTGCTGCTGCC 223427
QY 742 oAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyArgPheThrGluG1 762
Db 223426 TGGTGATCGTGTATCGAATGCATTGCCAGTGCAGCCAGGATTCGGTTTCCACCGAGG 223367
QY 762 ySerThrAspLysTyTyCysAlaTyArgGlyValTyLysProThrTyThrTh 782
Db 223366 CGAACCACTCAAGACCTTCTCTCGAGACATCACTGTGTGGGTCTCCACGTCCTCGTGGT 223307
QY 782 rGluTrpProAspCysAlaLysArgPheAlaAsnHisGlyPheLys---SerPheG1 801
Db 223306 G-----CCGACCTGCTGTCGGAGAACACCGAGCGCCCTTACACGTGACCGCTC 223253
QY 801 uMetPheTyLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAl 821
Db 223252 CATTACCTACCGCGCAATGGAGCAGTGGCCCAATCTGTCTGTGGTCTAGTACCAGGAGT 223193
QY 821 aPheGluThrThrLeuGly-----LysMetValProSerPheCysSerAsp-----836
Db 223192 GCTGGCACAGCACTATGGCGGACTCAACCCAGTTGCTCTCGCAGCGCTCTCGCGGTGAA 223133
QY 837 -----AlaGluAspIleAspCysArgLeuGluGluAsnLeuTh 849
Db 223132 TGTCAAATGAATGTGACCTTTGTGAAGTCTGTCCCATGCTGTGGAGGAATGTGGT 223073
QY 849 rLysLysTyCysLeuGluTyArgTyArgTyArgTyArgTyArgTyArgTyArgTyArg 869
Db 223072 CAAGATG-----223066
QY 869 yGlyTrpGlyAlaAlaAsnArgLeuAspTySerTyArgAspPheLeuAspThrValG1 889
Db 223065 -----GACTCATCTCTCCATTT--223048
QY 889 nGluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLe 909
Db 223047 -----CTGCCGCTGTGGTTCAGCGCGAGCTGTACGACT 223013
QY 909 uSerAspTyTyLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspG1 929
Db 223012 GTGCGGCTCCACGCTGAACCTGATCTTTGATCTG-----AGTGTACCTATGCGCAGTGC 222959
QY 929 u---ArgAsnAspThrLeuGluTrpGluAsnGlnArgLeuLeuGlnThrLeuGluTh 948
Db 222958 CGTCAATGATGACCTTTTGAACATTCGCAACATCGGTACCGAGTGTCTCGCTACGCGC 222899
QY 948 rIleThrAsnLysLeuLysArgThrLeuAsnLysAspProMetTySerPheGlnLeuAl 968
Db 222898 CCTCAAGTCCCAATCTCGCAGGATTTACTGCAT-----GT 222860
QY 968 aSerGluIleLeuIleAlaAspSerAsnSerLeuGluThrLysLysAlaSerProPheCy 988
```

Db	222859	GGGCGAGGTACTGAACATGACACACGAGCATGTGCGG	222823	Db	221926	CTGCCAGGCTTGGAGCAGGGTCGTACTCAAGTCGAGTGTGTGCCCGATGGGGGCAT	221867
Qy	988	sArgProGlySerValLeuArgGlyArgMetCysValAsnCysProLeuGlyThrTyrt	1008	Qy	1302	uHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsnAlaValCy	1322
Db	222822	-----CGTTGCGCTGCATCTGCGCGCGGAACGTAATGT	222791	Db	221866	CCACTGTGAGCAGAATATCAATGATGTTTCGGAGATCCCTGCTGTGGCGCCCACTG	221807
Qy	1008	rAsnLeuGluHisPheThrCysGluSerCysArgIleGlySerTyrglnAsnGluGlu	1028	Qy	1322	sGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArgCy	1342
Db	222790	GTCCAGGGGTGAGAACAGTCAGCTACTGCGCGAGGGGTACTACCAAGAACCGTGACCG	222731	Db	221806	CACAGATCTGGTCAATGATCTCCAGTGCCTGTCCGCCAGGATTTACGGGCAAGCGATG	221747
Qy	1028	YglnLeuGluCysValLeuCysProSerGlyMetTyThrGluTyrlleHisSerArgAs	1048	Qy	1342	sGlyLysAsnValAspGluCysLeuSerGlnProCysLysLysAsnGlyAlaThrCysLysAs	1362
Db	222730	CCAGGGAACCTGCGCTGCGCGCGCGAACCCTACCAAGAGAGGGGCCCAAGTC	222671	Db	221746	CGAGCAAAAGATCGATCTCTGCTATCGCAACCATCCAGCATGCG	221690
Qy	1048	nileSerAspCysLysAlaGlnCysLysGlnGlyThrTyrsertyrSerGlyLeuGluTh	1068	Qy	1362	pGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCysGluLe	1382
Db	222670	GCAGCGGACTGCATTCGCTGCGGTATGGACCTACTACCCACCGACTGTGTGC	222611	Db	221689	TCGCTGTTCGATCATCAGAGTGTGTTTGCATCCGGCTCGACGGGATCCGCTCGCACAT	221630
Qy	1068	rCysGluSerCysProLeuGlyThrTy	1085	Qy	1382	uAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGluLe	1402
Db	222610	GTGCTGGAGTGTGCGCGGAACTCATCTACTGCGGAACCAACACCGGTGGATTCAAGGA	222551	Db	221629	CAACATCGACGACTCGAGACCGACCCCTGCGCCATAGGGGAACCTGCGTCCGACTGGT	221570
Qy	1085	rCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaValAsnIleSerAl	1105	Qy	1402	uAsnSertyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu-ThrGluG	1422
Db	222550	TTGCCAGGCTGTCCGCGACAGAGCTTCACTTACCGCGCGGTGCTCGAAACAAGATCT	222491	Db	221569	CGACGGCTATAGTGCACACTGTGAACCCGCTACACGGCAAGATTCGACGACACCAT	221510
Qy	1105	aCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMetProCysHisPr	1125	Qy	1422	InSerThrGlyPheAsnLeuAspPheGluValserGlyIleTyrglyTyrrValMetLeuA	1442
Db	222490	GTGTGCGCCCAAGTGTGCGCGGAACTGCTACTGCGCACCGGACTGGCACCTGTGCGC	222431	Db	221509	CGACGACTGCG	221499
Qy	1125	oCysProArgAspTyrtyrGlnProAsnAlaGlyLysAlaPheCysLeuAlaCysProPh	1145	Qy	1442	sGlyMetLeuProSerLeuHisAlaLeuThrCys	1458
Db	222430	CTGCCCACTGCATCATTACACAGGAGCGCGGGTGGCAGAGTGCACAGATGTCGAG	222371	Db	221498	-----CCTCGAATCCCTGCAGCAGCGCGCCACCTGTGTGGACCGCTGGATGG	221450
Qy	1145	eTyrglyThrProPheAlaGlySerArgSerileThrGluCysSerSerPheSerSe	1165	Qy	1458	ySerSerAspAspMetAsnTyrglyThrProIleSerTyrrAlaValAsnGlySerA	1478
Db	222370	TAAATGAGAACCGATTACCGCGCTCCAAAGGAGCGCGAACAGTGC	222325	Db	221449	CTTCAGCTGCAATTCGCGCCCTGGGTACCTGGGTCTCTCTCTGCG	221406
Qy	1165	rThrPheSerAlaAlaGluSerValProAlaSerLeuGlyHisIleIleLysly	1185	Qy	1478	sPAsnThrLeuLeuLeuThrAspTyrrAsnGlyTyrrValLeuTyrrValAsnGlyArgGluL	1498
Db	222324	-----AAGCGGTGGTA	222313	Db	221405	-----AGCCGA	221399
Qy	1185	sArgHisGluIleSerSerGlnValPheHisGluCysPheAsnProCysHisAsnSe	1205	Qy	1498	yAlleThrAsnCysProSerValAsnAspGlyArgTyrrHisIleAlaIleThrTrpt	1518
Db	222312	-----TGTGTGAAGGTGCTTGCCAGCAGCG	222287	Db	221398	GATCGACGAGTGTCTGAGCAGCCCTGCAATCCGGTGGGCAACG	221355
Qy	1205	rGlyThrCysGlnLeuGlyArgGlyTyrrValCysLeuCysProLeuGlyTyrrThrGl	1225	Qy	1518	hrSerAlaAsnGlyIleTyrrLysValTyrrIleAspGlyLysLeuSerAspGlyAlaG	1538
Db	222286	CGACTGTGTGTCATGGGCCATGACATCCAGTGTCTGTGTCGCGCGGATTCTCTGG	222227	Db	221354	--AGCGTCCCTCGATCTGCACACAAAT	221312
Qy	1225	yLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeuAsnAsnGlyVa	1245	Qy	1538	lyLeuSer-ValGlyLeuProIleProGlyGly	1552
Db	222226	AGTGTGCTGGAACAGGACATGACAGGTGCGCTCCAGCCCTGCTCAATATGGTGTCA	222167	Db	221311	GGACGATTCAGAGGACCCCTGTGCGCCACGACATCGATGACTGCGAGCGCGCGCTG	221252
Qy	1245	lCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyrrThrGlyGlnAr	1265	Qy	1553	-Leu-----GlyGlnGluAsnLysLysGlyGlu	1567
Db	222166	GTGCAAGATCTGCGCGCAGGGCTATCGCTGTGAGTGTGCGCGGTGATATCTCGGCATCAA	222107	Db	221251	TCGTGAACACGGCATCTGTGCGGATCGCTCGTGTGCTTGTAGTGTGCGCGCAGCGG	221192
Qy	1265	sCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLysGlyIleCysVa	1285	Qy	1567	aGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrrAspTyrrValLeuSerProGl	1587
Db	222106	TTGCCAGGAGGAGCGCAGTGTGTGGCAACGACACCTGTCCGCGCGCGGCGCATGTGCAA	222047	Db	221191	A-----TGAGTGGCATGCGCTCGGACGA	221168
Qy	1285	lAsp-----GlyValAlaGly	1290	Qy	1587	nglnValLysSerLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAl	1607
Db	222046	GAACGACCGGGCTACAGAACGTGACCTGTGTGTCGCGCAGTGGCTACACCGGGATCA	221987	Db	221167	GCAGGTGACACCGTGGGAGCTCAGCGCGCTGCCAGACGATGCCAGCTCGACCT	221108
Qy	1290	-----	1290	Qy	1607	atrpProAspPheLeuSerGlyIleValGlyLysValIleAspSerLysSerIlePh	1627
Db	221986	GTGCGACGTGACCATCATCGGTGACGCGCGGAATGCGAATCCGTGCGAAACGAGCCAG	221927	Db	221107	GTTCAGGACTAC	221093
Qy	1291	-----TyrrArgCysThrCysValLysGlyPheValGlyLe	1302	Qy	1627	eCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAs	1647
				Db	221092	CTGGGTG---TGTCCAGC---GGCACCGATGGCAAGAACTGCGAGACCGCTCGGAA--	221041

```

QY 1647 pLeuLysProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAs 1667
Db 221040 -----CGTGCATCGGTGA 221027
QY 1667 nProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGlu-- 1686
Db 221026 TCCT-----TGCATGCACGCTGCAAG-----TGCACGGA 220997
QY 1687 -----ArgLysSerCysGlyValProProProLeuGluAsnGlyPheHisSerAl 1703
Db 220996 CTTTGGCTCTGGTCTTAACAGTGTGCCCT-----GC 220964
QY 1703 asAspPheTyrAlaGlySerThrValThrTyrGln----- 1715
Db 220963 GGAT-----TACTCGGCACATGGGTGTCTAGTACGAGTACACCGCATGCCAGGACATGT 220910
QY 1716 -CysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySe 1735
Db 220909 CTGTCAAGATGCG-----GCCACTTGTGTGCACAAATGGTGC 220874
QY 1735 rTrpAsnGlyValSerProSerCys-----LeuAs 1745
Db 220873 T-----GGCTACAGCTGCCAGTCCACCTTCCACCGTGCACAAATGCCAACACGGA 220820
QY 1745 pValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAs 1765
Db 220819 CATCGTGGAGTGCACGACAACTCT-----TGCCACCGCGGCCACGTCGTGTGATCTAAC 220763
QY 1765 pGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaG1 1785
Db 220762 CAACGCTTCTACTGTCTAGTGC-----CGCTTCAATATGACCGGAGACGATGCGCGCA 220709
QY 1785 uProLysLys 1788
Db 220708 GGCCATCCAA 220699

RESULT 23
LOCUS AX869682 801 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 4587 from Patent EP1074617.
ACCESSION AX869682
VERSION AX869682.1 GI:40024545
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 4587 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..801
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1 41e-76 Length: 801
Score: 1304.00 Matches: 247
Percent Similarity: 91.18% Conservative: 1
Best Local Similarity: 90.81% Mismatches: 16
Query Match: 12.76% Indels: 8
DB: 6 Gaps: 0

US-09-977-053-6 (1-1842) x AX869682 (1-801)
QY 318 TyrGluCysThrAlaCysProSerGlyThrTyrLysProGlySerProGlyGlyIle 337

```

```

Db 1 TATGAATGCACAGCTTGCCCATCGGGGACATACAAACCTGAAGCCTCCACCAGGAGGAATC 60
QY 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
Db 61 AGCAGTGTGCATTCATGTCCGATGAAATCACACCTCTCCCTCGGAGACATCCCT 120
QY 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
Db 121 GAAGACTGTGTCTGCAGAGAGGATACAGGCGCATCTGCCACAGACTGTGAACCTTGTCCAC 180
QY 378 CysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHis 397
Db 181 TGGCTCTCCCTCGAGCGCTCCGAAAATGGTTACTTTATCCAAAACACTTGCACACACAC 240
QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIle 417
Db 241 TTCATGTGCACCTGTGGGGTCCGATGTCACCTCGATTTGATCTTGTGGAGAGCAGCATC 300
QY 418 IleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThr 437
Db 301 ATCTATGTCTACCAATGTGTGTGCTCCGTTTCAGAGAGCTACTGCAGAGTAGAACA 360
QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetIleu 457
Db 361 TGTCTCTCATCTCCGCCACGCGAAAACATGCGGCACATCAGCTGTCTACAAAGGGAATGTTA 420
QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
Db 421 TATAAGACAACATGTGTGTGCTGTGTGTAAGGGTACAGACTAGAAAGCAGTGTATAG 480
QY 478 LeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHis 497
Db 481 CTTACTTGTCAAGGAACACAGCCAGTGGATGGCGCAGAACCCCGTGTGTGGAGCGCCAC 540
QY 498 CysSerThrPheGlnMet-ProLysAspValIleIleSerProHisAsnCys-GlyLysG 517
Db 541 TGTTCACCTTTTCAGATGCCCAANAGATGTCATCATATCCCCCAACACTGTGGGCAAGC 600
QY 517 InProAlaLysPheGlyThrIleCysTyrVal-SerCys-ArgGlnGlyPheIleLeuSe 536
Db 601 ANCCAGCCAAATTTGGGACGATCTGTATGTATAGTTGCCCGCCAGAGGTTTCATTTTATC 660
QY 536 rGlyValLysGluMetLeuArgCysThrThrSerGlyLys--TrpAsnValGlyValGln 555
Db 661 TGGAGTCANAGAAATGCTGANATGTACAACTCTCGGAAAATTTGGAATGTCGAGTTTCAA 720
QY 556 -AlaAlaVal-CysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluA 575
Db 721 GGCAGCTGTGTGTAAAGACNTGGAAGGNCCTCAATCAACTGGTCTCTAAGGNCATTANAG 780
QY 575 lalysThrLeuGluGlnGln 581
Db 781 GGTAAAGACTCGGGGAACAG 800

RESULT 24
LOCUS BD149744 801 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD149744
VERSION BD149744.1 GI:27855502
KEYWORDS JP 2002191363-A/4587.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 4587 09-JUL-2002;
HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
FN JP 2002191363-A/4587

```

PD 09-JUL-2002
PP 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
PT source 1. .801
PT Location/Qualifiers
1. .801 /organism='Homo sapiens (human)'.
FEATURES
source
1. .801
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.41e-76 Length: 801
Score: 1304.00 Matches: 247
Percent Similarity: 91.18% Conservative: 1
Best Local Similarity: 90.81% Mismatches: 16
Query Match: 12.76% Indels: 8
DB: 6 Gaps: 0
US-09-977-053-6 (1-1842) x BD149744 (1-801)
QY 318 TTTGluCysThrAlaCysProSerGlyThrTyrLysProGluCysSerProGlyGlyLe 337
DB 1 TATGAATGCACAGCTTGCCTCCATCGGGACATACAAACCTGAAGCCCTCACCAGGAGGAA 60
QY 338 SerSerCysIleProCysProAspGluAsnHisThrSerProGlySerThrSerPro 357
DB 61 AGCAGTTGATTCATGTCCTCCGATGAATACACCTCTCCACCTGGAGGACATCCCT 120
QY 358 GluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHis 377
DB 121 GAAGACTGTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGCCAC 180
QY 378 CysProAlaLeuLysProProGluAsnGlyTyrPheileGluAsnThrCysAsnAsnHis 397
DB 181 TGGCCCTGCCCTGAAGCTCCCGAAATGGTTACTTTATCCAAACACTTGCACACACCAC 240
QY 398 PheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerile 417
DB 241 TTCATGCGAGCCCTGGGGTCCGATGTCACTCCCTGGATTTGATCTTGTGGAGCAGCNC 300
QY 418 IleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThr 437
DB 301 ATCTTATGTCTACCCAATGGTTGTGGTCCGGTTCAGAGAGCTACTGCAGAGTAAGAACA 360
QY 438 CysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeu 457
DB 361 TGTCTCTCATCTCCGCCAGCGAACAATGGCCACATCAGCTGTGTCTACAAAGGAAATGTTA 420
QY 458 TyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLys 477
DB 421 TATAGACAAACATGTTTGTGCTGTGTATGAAGGGTACAGACTAGAGGCAGTATNAG 480
QY 478 LeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHis 497
DB 481 CTTACTTGTCAAGAAACACCAAGTGGATGGATGGCCAGAACCCCGGTGTGTGGAGCGCCAC 540
QY 498 CysSerThrPheGlnMet-ProLysAspValIleIleSerProHisAsnCys-GlyLysG 517
DB 541 TGTTCACCTTTCAGATGCCAANAGATGTCATATATCCCCCAACAACTGTGGGCAAGC 600
QY 517 InProAlaLysPheGlyThrIleCysTyrVal-SerCys-ArgGlnGlyPheIleLeuSe 536

DB 601 ANCCAGCCAAATTTGGGACGATCTGCTATGTTAAGTTGCCCGCAAGGGTTCATTTTATC 660
QY 536 rGlyValLysGluMetLeuArgCysThrThrSerGlyLys--TrpAsnValGlyValGln 555
DB 661 TGGAGTCAAGAAATGCTGATATGTACAACTTCTGGGAAATTTGGAATGTCGAGTTCAA 720
QY 556 -AlaAlaVal-CysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGluA 575
DB 721 GGCAGCTGTGTGAAGACGTCGAGAGNCCTCAATCACTGCTCTTAAGNCATTANAG 780
QY 575 laLysThrLeuGluGlnGln 581
DB 781 GGTTAAGACTCNGGGAACAG 800
RESULT 25
LOCUS AX540417 3262 bp DNA linear PAT 23-NOV-2002
DEFINITION Sequence 29 from Patent WO02055738.
ACCESSION AX540417
VERSION AX540417.1 GI:25273450
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Panzer, S.R., Lincoln, S.E., Altus, C.M., Dufour, G.E., Hillman, J.L.,
Jones, A.L., Dam, T.C., Liu, T.F., Harris, B., Flores, V., Daffo, A.,
Marwaha, R., Chen, A.J., Chang, S.C., Gerstin, B.H., Peralta, C.H.,
David, M.H. and Lewis, S.A.
TITLE Molecules for disease detection and treatment
JOURNAL Patent: WO 02055738-A 29 18-JUL-2002;
INCITE Genomics, Inc. (US)
FEATURES
source
1. .3262
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/notes="Incite ID No: LI:202943.4:2001JAN12"
ORIGIN
Alignment Scores:
Pred. No.: 2.43e-71 Length: 3262
Score: 1237.00 Matches: 240
Percent Similarity: 93.13% Conservative: 4
Best Local Similarity: 91.60% Mismatches: 9
Query Match: 12.10% Indels: 9
DB: 6 Gaps: 2
US-09-977-053-6 (1-1842) x AX540417 (1-3262)
QY 1590 LysSerLeuAlaThrSerCysProGluGluSerLysGlyAsnValLeuAlaTrrPro 1609
DB 289 GAGTCACTGGCTACCTCTGCCAGAGGAACCTAGTAAAGGAACCGTTAGCATGGCCT 348
QY 1610 AspPheLeuSerGlyLysValLysValLysValLysValLysValLysValLysVal 1629
DB 349 GATTCTTGTGAGGAATGTGGGGAAGTGAAGATCGATCTTAAGACATATTTGTCT 408
QY 1630 AspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLys 1649
DB 409 GGTGGCCACGCTTAGGGGGTTCAGTGCCTCATCTGAGAACTGCATCTGAAGATTTAAAG 468
QY 1650 ProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProVal 1669
DB 469 CCAGGTTCCAAAGTCAATCTGTCTGTGATCCAGGCTTCAGCTGTGGGAAACCTGTG 528
QY 1670 GlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnProLeuProHisCysGluArgIleSer 1689
DB 529 CAGTACTGTCTGAATCAGGACAGTGGACACCACTTCTCTCACTGTGAACGATTAGC 588
QY 1690 CysGlyValProProLeuGluAsnGlyPheHisSerAlaAspPheTyrAlaGly 1709

Db 589 TGTGGGTGTCACCTCTTTGGAGATGGCTTCCATTGAGCCGATGACTTCTATGCTGGC 648
Qy 1710 SerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe 1729
Db 649 AGCACAGTAACATACCAGTGCACAAATGGCTACTATCTATTGGGTGACTCAGGGATGTT 708
Qy 1730 CysThrAspAsnGlySer-TrpAsnGlyValSerProSerCysLeuAspValAspGluCy 1749
Db 709 TGTACAGTAAATGGAGCTAGGAACGGGCTTTACCATCTCTGCCGTGATGTCGATGAGTG 768
Qy 1749 sAlaValGlySerAspCysSerGluHisAlaSerCysLeu-AsnValAspGlySerTyrI 1769
Db 769 TGCAGTTGGATCAGATTGTAGTACGATGCTTTCTTGCTGAACGCTAGATGCTCTACA 828
Qy 1769 leCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProIleLysC 1789
Db 829 TATGTTTCATGTGTCCTCCACCGTACACAGAGATGGGAAAACTGTGCGAACCCTATAAAT 888
Qy 1789 ys-LysAlaProGly---AsnProGluAsnGlyHisSerSerGlyGluIleTyrThr-Va 1807
Db 889 GCTAAGGCTCCAGCGCAGATCCGGAAATGSCCACTCTCAGGTGAGATTATACAGT 948
Qy 1807 lGlyAla---AlaValThrPheSerCysGlnGluGlyTyr-GlnLeuMetGlyValThrL 1826
Db 949 AGGTGCCCGAAGTCACATTATTAGGTGTACGAGAGATACCAGTTGATGGGAGTAACCA 1008
Qy 1826 ysIleThrCysLeuGlu-SerGlyGluTrpAsnHisIleLeuPro-TyrCys 1842
Db 1009 AAATCACAATGTTGGAGTACTGGAGAAATGGAATCATCTAATACCAATATTGT 1060

Search completed: May 10, 2004, 06:20:52
Job time : 15424 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 9, 2004, 12:50:18 ; Search time 27508.5 Seconds
(without alignments)
5626.549 Million cell updates/sec

Title: US-09-977-053-4
Perfect score: 19973
Sequence: 1 MWPLAFCCWGLALVSGWAT.....CHCLSSWTGHCNSRKRTGTF 3571

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:
-MODEL=frame+_{p2n}.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US09977053/runat_06052004_075942_18188/app_query.fasta_1.5710
-DB=GenEmbl -START=1 -END=1 -MATRIX=blosum62 -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -FWT=1 -RND=1 -TRANS=human40.cdi -LIST=100
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=25 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977053 -CGCN 1.1 27696 @runat_06052004_075942_18188 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pt:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_nam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19958	99.9	10878	6	AR435509 Sequence
2	19958	99.9	10878	6	AX375161 Sequence
3	19616	98.2	11152	6	AX686198 Sequence
4	19566	98.0	11158	6	AX686196 Sequence
5	16736	83.8	11289	10	AF206329 Mus muscu
6	16707	83.6	11230	6	AR435510 Sequence
7	16707	83.6	11230	6	AX375163 Sequence
8	9400	47.1	5421	9	BD127949 Primer fo
9	9400	47.1	5421	9	AK075235
10	8427	42.2	6153	6	AX350821 Sequence
11	8157	40.8	4701	9	BSX537918 Homo sapi
12	7537	37.7	5124	6	AX880905 Sequence
13	7537	37.7	5124	6	BD158659 Primer fo
14	7537	37.7	5124	9	AK027870 Homo sapi
15	7486	37.5	4385	9	BSX538049 Homo sapi
16	7471	37.4	3991	6	AX367105 Sequence
17	6950	34.8	4088	6	BD127948 Primer fo
18	6950	34.8	4088	9	AK075234 Homo sapi
19	6593	33.5	3914	9	AL832416 Homo sapi
20	6218	31.1	194835	9	AL158158 Human DNA
21	6044	30.3	3706	6	AX050034 Sequence
22	5142.5	25.7	26868	2	AC096906 Rattus no
23	5086.5	25.5	152909	10	AL929406 Mouse DNA
24	4930	24.7	3262	6	AX540417 Sequence
25	4776	23.9	3253	9	AK122605 Homo sapi
26	4674	23.4	2575	6	AX430405 Sequence
27	4376.5	21.9	167023	2	BSX548016 Danio rer
28	4039.5	20.2	2450	9	HST000009
29	3596	18.0	1952	6	AX350819 Sequence
30	3177.5	15.9	2388	10	BC008135 Mus muscu
31	3073	15.4	1969	6	AX880735 Sequence
32	3073	15.4	1969	6	BD158559 Primer fo
33	3073	15.4	1969	9	AK023591 Homo sapi
34	2831	14.2	1765	6	BD127875 Primer fo
35	2831	14.2	1765	9	AK075200 Homo sapi
36	2638	13.2	1653	9	AF308289 Homo sapi
37	2622	13.1	1747	6	AX013771 Sequence
38	2622	13.1	1747	6	BD205162 Human nuc
39	2506	12.5	1357	6	AX430265 Sequence
40	2374	11.9	1760	6	BD229966 Method an
41	2305	11.5	1709	6	AX350817 Sequence
42	2284	11.4	1766	6	AX050032 Sequence
43	2282	11.4	1638	6	I76197 Sequence 1
44	2248	11.3	1377	6	AX050029 Sequence
45	2038	10.2	6044	9	CHPCR1WT L24920 Chimpanzee
46	2032	10.2	7313	6	AR338768 Sequence
47	2027	10.1	6951	6	AR380720 Sequence
48	2027	10.1	6951	6	AX774977 Sequence
49	2027	10.1	6951	9	Y00816 Human mRNA
50	2016	10.1	6951	6	AR364181 Sequence

51	2015	10.1	6000	9	BABCR1A	L39791 Papio hamad
52	1875.5	9.4	11580	9	AF333704	AF333704 Homo sapi
53	1852	9.3	12525	9	AX374893	AX374893 Sequence
54	1836.5	9.2	11221	10	AY017475	AY017475 Mus muscu
55	1825.5	9.1	10433	6	AX374896	AX374896 Sequence
56	1825.5	9.1	10673	6	AX374891	AX374891 Sequence
57	1822	9.0	10136	6	AX685957	AX685957 Sequence
58	1792.5	9.0	10774	9	AB067481	AB067481 Homo sapi
59	1784	8.9	13040	9	AB114605	AB114605 Homo sapi
60	1784	8.9	13148	9	AB114604	AB114604 Homo sapi
61	1774	8.9	9120	10	AK122567	AK122567 Mus muscu
62	1729	8.7	12486	9	AY210419	AY210419 Homo sapi
63	1701.5	8.5	13113	9	AY210418	AY210418 Homo sapi
64	1694.5	8.5	109766	2	AC117712	AC117712 Mus muscu
65	1614	8.1	10944	9	AY017307	AY017307 Homo sapi
66	1606.5	8.0	5420	6	AX333690	AX333690 Sequence
67	1606.5	8.0	5420	6	HSCR1R	X05309 Human mRNA
68	1604.5	8.0	5420	6	AR364182	AR364182 Sequence
69	1465.5	7.3	109766	2	AC117712	AC117712 Mus muscu
70	1447	7.2	765	6	AX870332	AX870332 Sequence
71	1447	7.2	765	6	BD150394	BD150394 Primer fo
72	1443.5	7.2	18583	2	AC018297	AC018297 Drosophil
73	1443.5	7.2	148780	3	AC008326	AC008326 Drosophil
74	1443.5	7.2	174287	3	AC007977	AC007977 Drosophil
75	1443.5	7.2	270766	3	AE003615	AE003615 Drosophil
76	1379	6.9	3896	6	AX149475	AX149475 Sequence
77	1349	6.8	3905	6	AX149473	AX149473 Sequence
78	1341	6.7	706	6	AX677660	AX677660 Sequence
79	1304	6.5	801	6	AX869682	AX869682 Sequence
80	1304	6.5	801	6	BD149744	BD149744 Primer fo
81	1225.5	6.1	771	6	BD126501	BD126501 Primer fo
82	1217	6.1	4094	9	HUMEB2CR2	J03565 Human Epste
83	1209.5	6.1	3518	9	AK127129	AK127129 Homo sapi
84	1184	5.9	3934	6	AR380521	AR380521 Sequence
85	1184	5.9	3934	9	HUMEBVR	M26004 Human CR2/C
86	1182	5.9	8010	6	AX685959	AX685959 Sequence
87	1173.5	5.9	728	6	BD125275	BD125275 Primer fo
88	1170.5	5.9	728	6	BD126535	BD126535 Primer fo
89	1170.5	5.9	3923	6	AX774975	AX774975 Sequence
90	1170.5	5.9	3923	9	HSBLCR2	Y00649 Human mRNA
91	1126.5	5.6	3187	10	MUSCR2B	M61132 Mouse compl
92	1124.5	5.6	3220	10	MUSCR2AA	M35684 Murine comp
93	1124.5	5.6	5463	9	AB067471	AB067471 Homo sapi
94	1117.5	5.6	251269	2	AC145193	AC145193 Gallus ga
95	1114.5	5.6	4020	4	AF038131	AF038131 Ovis arie
96	1086.5	5.4	761	6	BD125309	BD125309 Primer fo
97	1060	5.3	594	6	BD125308	BD125308 Primer fo
98	1059	5.3	4849	3	AK113070	AK113070 Ciona int
99	1058.5	5.3	6004	6	AX537467	AX537467 Sequence
100	1045.5	5.2	4003	4	SSC278470	AJ278470 Sus scrof

ALIGNMENTS

RESULT 1
AR435509 AR435509 10878 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 1 from patent US 6656707.
DEFINITION
ACCESSION AR435509
VERSION AR435509.1 GI:40198412
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 10878)
AUTHORS Welch, A. A. and Elliott, G. S.
TITLE C3b/C4b complement receptor-like molecules and uses thereof
JOURNAL Patent: US 6656707-A 1 02-DEC-2003;
FEATURES
Location/Qualifiers
1..10878
/organism="unknown"
/mol_type="genomic DNA"

Alignment Scores:

Pred. No.:	0	Length:	10878
Score:	1958.00	Matches:	3566
Percent Similarity:	99.94%	Conservative:	3
Best Local Similarity:	99.8%	Mismatches:	2
Query Match:	99.92%	Indels:	0
DB:	6	Gaps:	0
US-09-977-053-4 (1-3571) x AR435509 (1-10878)			
QY	1	MetTTPProAqGLeuAlaPheCysCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr	20
DB	11	AVGTGGCCCTGGCTGGCCCTTTTGTGTGGGTCTGGCGCTGGTGGGGAGC	70
QY	21	PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro	40
DB	71	TTTCAGCAGATGTCCCGTGGCGCAATTTACGTTCCGCTCTTCCCGAGCCGGCCC	130
QY	41	GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg	60
DB	131	GGGGCCCCGGGAGTATCCCGCGCGCGCTCTTGGCGAGCAGAGCGCGGGAGCAGA	190
QY	61	ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu	80
DB	191	GTGGAGCGGCTGGGCGCGCGGTTCGGCGACCGGTCTGGCGGAGCTCAGCGAG	250
QY	81	ArgLeuGluLeuValPheLeuValAspSerSerSerValGlyGluValAsnPheArg	100
DB	251	CGCTGGAGCTGTCTTCTGTGTGATGTTCTCCAGGTGGCGGAGTCACTTCCTCGC	310
QY	101	SerGluLeuMetPheValArgIleLeuSerAspPheProValValProThrAlaThr	120
DB	311	AGCGAGCTCATGTTCTGGTCCGCAAGCTGTCTCCGCTTCCTCCGTTGGTCCCGCACG	370
QY	121	ArgValAlaIleValThrPheSerSerIleValProArgValAspTyrIle	140
DB	371	CGCTGGCGCATGTCGACCTTCTGTCGAAGAATGAGTGTGGCGCGCTCGATTACATC	430
QY	141	SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnIleProAlaIle	160
DB	431	TCCACCGCGCGCGCGCGCAGCAGTCTCTCCAGAGATCCCTGCGCATC	490
QY	161	SerTyrArgGlyGlyGlyThrTyrThrIleGlyAlaPheGlnGlnAlaGlnIleLeu	180
DB	491	TCCTACCGAGGTGGCGGCACCTACACAGGCGCGCTTCAGCAAGCCCGCAATTCCT	550
QY	181	LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn	200
DB	551	CTTCATGCTAGAGAAACTCAACAAAGTTGTATTCTCATCTGATGGATATTCAT	610
QY	201	GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr	220
DB	611	GGGGAGACCCCTAGACCAATTCAGCGTCACTGCGAGATTCAGGAGTGGAGATCTTCACT	670
QY	221	PheGlyIleTTPGlnGlnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu	240
DB	671	TTTGGCAATGCGAAGGGAACATTCAGAGCTGAATGATGATGCTCCACCCCAAGGAG	730
QY	241	GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu	260
DB	731	GAGCATGTGTACCTGCTACACAGTTTTCAGATTTAGGCTTTAGCTCGCGGGCATTTG	790
QY	261	HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu	280
DB	791	CATGAAGATCTACCTCTGGGAGTTTATTCAAGATGATATGCTCCACTGCTCTTATCTT	850
QY	281	CysAspGluGlyLysAspCysCysAspArgMetGlySerCysGlyThrHisThr	300
DB	851	TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGAGAGCTGCAATGTGGGACACACA	910
QY	301	GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys	320

911 GGCCATTTTTCAGTCATCTCTGAAAGGGGTATTACGGGAAGGCTCTGCAGTATGAATGC 970
321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyLysSerSerCys 340
971 ACAGCTTGGCCATCGGGGACATACAAACCTTGAAGGCTCACAGGAGGAATCAGCAGTGC 1030
341 IleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAspCys 360
1031 ATTCCATGTCCTGATGAAATATCACCTCTCCACCTGGAGACATCCCTCGAAGACTGT 1090
361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
1091 GTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGAACCTGTGTCCACTGCCCTGCC 1150
381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
1151 CTGAAGCCTCCCGAAATGGTTACTTTATCCAAACACATTGCAACCACTTCAATGCA 1210
401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
1211 GCCTGTGGGGTCCGATGTACCCCTGGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1270
421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
1271 CTACCCATGGTTTGTGTGGTCCGTTCCAGAGACTACTGACAGTAAAGACATCTCTCAT 1330
441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
1331 CTCCGCGAGCCGAAACATGGCCACATCAGCTGTTCTACAGGGGAAATGTTATATAAGACA 1390
461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
1391 ACATGTTTGTGTCCTGTGTGATGAAGGTAACAGACTAGAGGAGTGTGAAGCTTACTTGT 1450
481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
1451 CAAGGAACAGCCAGTGGGATGGCCAGAACCCGGTGTGTGGAGCGCACTGTTCACC 1510
501 PheGlnMetProLysAspValIleSerProHisAsnCysGlyLysGlnProAlaLys 520
1511 TTTTCAGATGCCAAAGATATCATCATATATCCCCACAACTGTGGCAGCAGCCAGCNA 1570
521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
1571 TTTGGAGCATCTGCTATGATGATGATGTCGCCCAAGGGTTCATTTTATCTGGAGTCAAGAA 1630
541 MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaValCysLys 560
1631 ATGCTGAGATGTACCACTTCTGGAAATGGAATGTCCGAGTTCAGGCAGCTGTGTGTA 1690
561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
1691 GAGCTGGAGGCTCTCCAAATCACTGTCTCTAAGGACATAGAGGCTAAGACTCTGGAACAG 1750
581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyLys 600
1751 CAAGATCTCGCCAAATGTATTCTGGCAGATTCACAGCTTAAGACAACTCTGTGTGAAAG 1810
601 ValSerValHisValHisProAlaPheThrProTyrLeuPheProIleGlyAspVal 620
1811 GTGTCACTCCAGTTCATCCAGCTTTCACCCCACTTACCTTTTCCAGTTGGAGATGT 1870
621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
1871 GCTATCGTATACAGCGCAACTGACCTATTCGGCAACAGCCAGCCAGCTGCAATTTTCCATATC 1930
641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
1931 AAGGTTATTGATGACAGAACCACTGTCTATAGACTGTGTGAGATCTCACTCCCTCCAG 1990
661 ValSerGluLysValHisAlaIleSerTrpAspGluProGlnPheSerAspAsnSerGly 680
1991 GTCTCGAGAGGTATACGCGCAAGCTGGATGAGCCTCAGTTCTCAGACAACTCAGGG 2050

QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyThr 700
Db 2051 GCTGAATGTGTCTATACCAGAGTCTATACAAAGGAGACCTTTTCCCTCAAGGGAGACT 2110
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Db 2111 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCATATT 2170
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
Db 2171 GTCATAAAAGGTTCTCCCTGTGAAATTCATTCACCTGTAAATGGGATTTTATATGC 2230
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
Db 2231 ACTCCAGATTAATACTGGAGTCAACTGTACATTAACCTTGTGGAGGCTATGATTTTCA 2290
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
Db 2291 GAAGGTCTACTGACAGTATTAATGTGCTTATGAGATGGCGTCTGGAAACCAACATAT 2350
QY 781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Db 2351 ACCACTGAATGGCCAGACTGTGCCAAAACAGTTTGTCTAAACACAGGCTTCAAGTCTTT 2410
QY 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
Db 2411 GAGATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAAGTTTTCTGNA 2470
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
Db 2471 GCATTTGAGAGCACCCTGGGAAAATGGTCCCATCATTTTGTAGTGATGACAGGACATT 2530
QY 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
Db 2531 GACTGCAGACTGGAGAGAACTGACCAAAAATATTTCCCTAGAAATAATAATATGACTAT 2590
QY 861 GluAsnGlyPheAlaIleGlyProGlyLysTrpGlyAlaAlaAsnArgLeuAspTyrSer 880
Db 2591 GAAAATGGCTTTGCAATTTGACACAGGTGGCTGGGTGACGTAATAGCTGGATCTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
Db 2651 TACGATGACTTCTGACACACTGTGCAAGAACAGCCACAGCATCGGCAATGCCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
Db 2711 TCACGGATTAAGAAAGAGTGCCTTATCTGACTATATAAATAAGTTTAAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTrpGluAsnGlnGln 940
Db 2771 ACAGCTAGTGTGCATTTACCCGATGAAGAANAATGATACCTTTGANTGGGAANAATCAGCA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
Db 2831 CGACTCTTCAGACATTTGAAACTATCACAAATAAACTGAAAGAGGACTCTCAACAAAGAC 2890
QY 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
Db 2891 CCCATGTATTCTCTTCAGCTTGCATCAGAAATAACTTATAGCCGACAGCAATTCATTAGAA 2950
QY 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
Db 2951 ACAAAGAAGCTTTCCCTCTCTCGACACAGGCTCAGTGTCTGAGAGGCGCTATGTGTGTC 3010
QY 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
Db 3011 AATTGCCCTTTGGAACTTATTAATCTCGAACATTTTCACTGTGAAAGCTGCGGATC 3070
QY 1021 GlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
Db 3071 GGATCTCTATCAAGTGAAGAAGGCACTTGGAGTGCAGCTTTTGCCTCTCTGGATGTAC 3130

Qy	1041	ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr	1060	Db	4211	GAATTAAATTCATACAGTTGTAAATGTCCAGCAGGATTTTCAGGCAAAAGGTTGAAACA	4270
Db	3131	ACGGAAATATATCCATTCAGAAACATCTCTGATTGTAAAGCTCAGTGTAAACAAGGCACC	3190	Qy	1421	GluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMet	1440
Qy	1061	TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys	1080	Db	4271	GAACAGTCTACAGGCTTTAACTGGATTTTGAAGTTTCTGGCATCTATGGATATGTCTAG	4330
Db	3191	TACTCATACATGAGCTTGAACCTTGTGAATCGTGTCCACTGGGCACTTATCAGCCAAA	3250	Qy	1441	LeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTrpMetLysSerSer	1460
Qy	1081	PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyVala	1100	Db	4331	CTAGATGGCATGCTCCCATCTCTCCATGCTCTAACTGTACCTTCTGGATGAAATCTCTCT	4390
Db	3251	TTTGGTTCCTCCGAGCTGCTCTCTGTGTCCAGAAACACTCACTGTGAAAAGAGAGCC	3310	Qy	1461	AspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThr	1480
Qy	1101	ValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeu	1120	Db	4391	GACGACATGAATATGGAACACCAATCTCTATGCACTGTGATTAACCGGACGACATACC	4450
Db	3311	GTGAACATTTCTGCATGTGGAGTTCCTGTGTCCAGAAAGGAAATTTCTCGCGTTCTGGGTTA	3370	Qy	1481	LeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyValArgGluLysIleThr	1500
Qy	1121	MetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCys	1140	Db	4451	TTGTCTCTGACTGATTAACCGCTGGGTCTTTATGTGAATGGCAGGGAAGAATACAA	4510
Db	3371	ATGCCCTGTCAACCATGCTCTCGTGACTATTACCAACCTAATGAGGGAAGGCCCTTCTGC	3430	Qy	1501	AsnCysProSerValAsnAspGlyValGtrPheHisIleAlaIleThrTrpThrSerAla	1520
Qy	1141	LeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCys	1160	Db	4511	AACTGTCTCCGTGGTGAATGATGGCAGATGGCATCATATTGCAATCACTTGGACAAGTGCC	4570
Db	3431	CTGGCCTGTCCCTTTATGGAACATACCCCATTCGCTGCTCCAGATCCATCACGAATGT	3490	Qy	1521	AsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSer	1540
Qy	1161	SerSerPheSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeu	1180	Db	4571	AATGGCATCTCGGAAGTCTATATCGATGGGAATTTATCTGACGGTGGTCTGCTCTCT	4630
Db	3491	TCAAGTTTTAGTTCAACTTTCTCAGCGGCAGAGGAAGTGTGTGCCCCCTGCTCTCT	3550	Qy	1541	ValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluGlnAspLysLys	1560
Qy	1181	GlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePheAsn	1200	Db	4631	GTTGGTTGGCCCATCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	4690
Db	3551	GGAATATATAAAGAGGAGCATGAATACAGAGTACAGTGTTCATGATGAATGCTTCTTTAAC	3610	Qy	1561	GlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrp	1580
Qy	1201	ProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCysPro	1220	Db	4691	GGAGAGGATTCAGCCAGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGG	4750
Db	3611	CTTGGCCACATAGTGGAACTGCCAGCAACTGGGCGTGTATGTCTCTGTCTCA	3670	Qy	1581	AspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluLeu	1600
Qy	1221	LeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCys	1240	Db	4751	GACTATGTCTCTTCCACAGCAGGTGAAGTCACTGGCTACTCTCTGCCAGAGAACTC	4810
Db	3671	CTTGGATATACAGGCTTAAAGTGTGAACACAGATCGATGATGTCAGCCACCTGCTTGC	3730	Qy	1601	SerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysValLys	1620
Qy	1241	LeuAsnGlnGlyValCysLysAspLeuValGlyLupheIleCysGluCysProSerGly	1260	Db	4811	AGTAAGGAACAGTGTAGCATGCGCTGATTTCTTGTTCAGGAATGTGGGGAAGTGAAG	4870
Db	3731	CTCAACATGAGGTGTGAAGACTAGTGTGGGAATTCATTTGTGAGTGCCCATCAGGT	3790	Qy	1621	IleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHis	1640
Qy	1261	TyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeuAsn	1280	Db	4871	ATCGATTTAAGACATATTTTGTCTGATTTGCCACGCTTAGGAGGCTCAGTGGCTCAT	4930
Db	3791	TACACAGTTCAGCGGTGTGAAGAAAAATATAATCAGTGTAGTCTCAGTCTCTGTTTAAAT	3850	Qy	1641	LeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspPro	1660
Qy	1281	LysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheVal	1300	Db	4931	CTGAGAACTGCATCTGAAGATTTAAAGCCAGGTTCCAAAGTCAATCTGTTCTGTATCCA	4990
Db	3851	AAAGGAATCTGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3910	Qy	1661	GlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThrGln	1680
Qy	1301	GlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAla	1320	Db	4991	GECTTCCAGCTGGTGGGAAACCCCTGTGTCAGTACTGTCTGTAATCAAGGACAGTGGACAA	5050
Db	3911	GECCTGCATTTGTGAACAGAACTCAATGAATGCCAGTCCAAACCCATGCTTAAATAATGCA	3970	Qy	1681	ProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPhe	1700
Qy	1321	ValCysGluAspGlnValGlyGlyPheLeuCysLysCysProProGlyPheLeuGlyThr	1340	Db	5051	CCACTTCTCTCATCTGGAACGCAATTTGTTCTGATTTGCCACGCTCTTGGAGAAATGGCTTC	5110
Db	3971	GTCTGTGAAGACCAAGTGTGGGGAATCTGTGTGCAAAATGCCACCTGGATTTTGGGTACC	4030	Qy	1701	HisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr	1720
Qy	1341	ArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCys	1360	Db	5111	CATTGAGCCGATGATCTTATGCTGGCAGCAGTNACTACCACTGACGACACATGGCTAC	5170
Db	4031	CGATGTGAAGAAAGACGTGCATGAGTGTCTCAGTCAAGCATATGCAAAATGGAGCTACCTGT	4090	Qy	1721	TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer	1740
Qy	1361	LysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHisCys	1380	Db	5171	TATCTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGAGCTGGAAACGGCTTCCA	5230
Db	4091	AAAGACGTGTCAATAGCTTCAGATGCTGTGTGAGCTGGCTTCACAGGATCACACTGT	4150	Qy	1741	ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer	1760
Qy	1381	GluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAsp	1400	Db	5231	CCATCTCTGCTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTG	5290
Db	4151	GAATTGAACATCAATGAATGTGCTGCTAAATCCATGTAGAAATCAGGCCACCTGTGTGGAT	4210	Qy	1761	CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly	1780
Qy	1401	GluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThr	1420				

Db 5291 TGCTGAACGTAGATGATCTACATATGTTTCATGTGTCCACCGTACACGAGATGG 5350
Qy 1781 LysAsnCyAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
Db 5351 AAAAACTGTGCAGAACCTATAAAATGTAAGGCTCCAGGAAATCCGAAAAATGCCACTCC 5410
Qy 1801 SerGlyGluIleThrValGlyAlaAlaValThrPheSerCysGlnGluGlyThrGln 1820
Db 5411 TCAGGTGAGATTTATACAGTAGTAGTGCGAAGTCACATTTTCGTGTGAGGAGATACCAG 5470
Qy 1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
Db 5471 TTGATGGAGTAGTACCAAAATCACATGTTTGGAGTCTGGAGATGGAATCATCTAATACCA 5530
Qy 1841 TyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
Db 5531 TATTTGTAAGAGCTGTTTCATGTGGTAAACCGGCTATTTCGAAAAATGGTTGCAATTGAGGAG 5590
Qy 1861 LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyThrLeuAla 1880
Db 5591 TTAGCATTTTACTTTTGGCAGCAAGTGACATATAGGTGTAAATAAGATATATCTTGACC 5650
Qy 1881 GlyAspLysGluSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCys 1900
Db 5651 GGTGATAAAGATCATCTCTGTGTCTGTAAAGTTCTTGAGTCAATCCCTCCTGTGTGT 5710
Qy 1901 GluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTrpIleLeuSerGly 1920
Db 5711 GAACCAAGTGAAGTGTCTAGTCCGAAAAATATAAATAATCGAAAAATATATTTTGAGTGGG 5770
Qy 1921 LeuThrTyrLeuSerThrAlaSerThrLysCysAspThrGlyThrSerLeuGlnGlyPro 1940
Db 5771 CTTACCTACTTCTTACTGCACTCATATTCATGCGAATACAGATACAGCTTACAGGCGCCCT 5830
Qy 1941 SerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeu 1960
Db 5831 TCCATTATTGAATGTCACGGCTCTGCGCATCTGGCAGACAGCGCCACCTGCTGTCTACCTC 5890
Qy 1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
Db 5891 GTCTTCTGTGGAGAACCTCGCATCAAGATGCTGTCAATACGGGGAATACTTCACT 5950
Qy 1981 PheArgAsnThrValThrTyrThrCysLysGluGlyThrThrLeuAlaGlyLeuAspThr 2000
Db 5951 TTCAGAACACCTGCTACTTACACTTCGAAAGAAGGCTATACTCTTCTGCTGTGACACC 6010
Qy 2001 IleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSer 2020
Db 6011 ATTGAATGCTGGCCGACGGCAAGTGGAGTAGAAGTAGACCAAGTGCCTGGGCTGTCTCC 6070
Qy 2021 CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly 2040
Db 6071 TGTGATGAGCCACCAATTTGGACCAACCGCTCTCCAGAGACTGCCCATCGGCTCTTTGGA 6130
Qy 2041 AspIleAlaPheTyrThrCysSerAspGlyThrSerLeuAlaAspAsnSerGlnLeuLeu 2060
Db 6131 GACATTTGCAATTTCTACTTACTTCTGTGTTTACGCTTACGCTACGACAAATTTCCAGCTTCTC 6190
Qy 2061 CysAsnAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAla 2080
Db 6191 TGCATGTCGCCAGGCAAGTGGGTACCCCCAGAGAGGTCAAGACATGCCCTGGTGTATAGCT 6250
Qy 2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
Db 6251 CATTTCTGTGAAAAACCTCCATCGGTTTCTATAGCATCTTGGAAATCTGTGAGCAAGCA 6310
Qy 2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
Db 6311 AAATTTGCAAGTGGCTCAGTTGAGCTTTAATGATGATGAGGCGCTTTGACTGAACACC 6370
Qy 2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
Db 6371 TCAGCAAGATTTGAATGTATGAGAGGTGGCGAGTGGAAACCTTTCCCCCATGTCTCATCCAG 6430

Qy 2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
Db 6431 TGCATCCCTGTGCGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCA 6490
Qy 2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
Db 6491 AACTACAGTTTTGAGGCCATGGTGGCTTACAGCTGCAACCAAGGGTCTTACATCAAGGG 6550
Qy 2181 GluLysLysSerThrCysGluAlaIleThrGlyGlnTrpSerSerProIleProThrCysHis 2200
Db 6551 GAAAGAAGACGACCTGCCAAGCCACAGGCGAGTGGAGTAGTCTCTATACCGACGTGCCAC 6610
Qy 2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
Db 6611 CCGTATCTTGTGTGTGAACCACTTAAGTTTGAGATGGCTTTCTGAGGACATACAACTGGC 6670
Qy 2221 ArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySer 2240
Db 6671 AGCATCTTTGAGAGTGAAGTGAGGTATCACTGTAAACCGGGCTATAAGTCAGTCGGAAGT 6730
Qy 2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
Db 6731 CCTGTATTTGTGTGCCAAGCAATCGCCACTGGCACAGTAGTAATCCCTCTGTATGTGT 6790
Qy 2261 ProLeuAspCysGlyLysProProIleGluAsnGlyPheMetLysGlyGluAsnPhe 2280
Db 6791 CCTCTGACGTGTGGAACCTCCCGATCCAGATGGCTTCATGAAGGAGAAACTTT 6850
Qy 2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
Db 6851 GAAGTAGGTTCCAAAGTTCAGTTTTTCTGTAAAGAGGTTATGAGCTGTGTGGTGACAGT 6910
Qy 2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320
Db 6911 TCTTGGACATGTGCAGAAATCTGGCAAATGGAATAAGAAAGTCAATCCAAAGTGTATGCT 6970
Qy 2321 AlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThrThr 2340
Db 6971 GCCAGTGGCCAGAGCGCCCTCTTGGAAAAACAGCTAGTATTAAGAGAGTTCACACC 7030
Qy 2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
Db 7031 GAGGTAGGAGTGTGACATTTCTCTGTAAAGAGGCGATGCTCTGCAAGGCCCCCTCTGTC 7090
Qy 2361 LeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
Db 7091 CTGAAATGCTTGGCATCCCAAGCAATGGAATGACTCTTTCTCTGTGTGTAAGATTGTCTT 7150
Qy 2381 CysThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
Db 7151 TGTACCCCACTCCCTAATTTCTTGTGTGTGCCATCTCTTCTTCTGCTCTCTCAATTT 7210
Qy 2401 GlySerThrValLysThrSerCysValGlyPhePheLeuArgGlyAsnSerThrThr 2420
Db 7211 GGAAGTACTGTCAAGTATCTCTGTGTAGTGGTGTCTTCTTCTTCTTCTTCTTCTTCTT 7270
Qy 2421 LeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys 2440
Db 7271 CTCTGCAACCTCTGATGGCACCTGAGCTCTCCACTGCCAGAAATGTGTTCAGTAGAATGT 7330
Qy 2441 ProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeuSer 2460
Db 7331 CCCCACTGTAGGAATCCCAATGGAATGTGTGTGTGAGGCTTGTGCTATCTCAGC 7390
Qy 2461 ThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeuCys 2480
Db 7391 ACAGCTCTCTATACCTGCAAGCCAGGCTTTGAATTTGTGTGTGAGGAAATCTACCCACTTGT 7450
Qy 2481 GlyLysGlnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLys 2500
Db 7451 GGAGAAAAATGTCATGCTGTGGAGGAAAAACCAACATGTAAGGCCATTTGATGCTGCTGAAA 7510

QY 2501 ProLysGluIleLeuAsnGlyLysPheSerThrAspLeuHisThrVal 2520
DB 7511 CCCAAGGAGATTTTGAATGGCAAAATCTCTTACACGACCTACACTATGGACAGACCGTT 7570
QY 2521 ThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu 2540
DB 7571 ACCTACTCTTGCAACCGAGGCTTTGGCTCGAAGGTCCAGTGCCTTGACCTGTTTAGAG 7630
QY 2541 ThrGlyAspTTPAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln 2560
DB 7631 ACAGGTGATTTGGATGTAGATGCCCATCTTGCAATGCCATCCACTGTGTATCCCCACAA 7690
QY 2561 ProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyr 2580
DB 7691 CCCATTTGAAAATGGTTTGTAGAGGTGCAGATTACAGCTATGGTGCCATATCATCTAC 7750
QY 2581 SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly 2600
DB 7751 AGTTGCTTCCCTGGGTTCAGGTGGCTGTGTATGCCATGCAGACCTGTGGAAGGTACAGA 7810
QY 2601 TrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp 2620
DB 7811 TGGTCAAGTTCCATCCCAACATGATATGCCAATAGACTGTGGCTCTCCCTCATATAGAT 7870
QY 2621 PheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMet 2640
DB 7871 TTTGGAGACTGTACTAAACTCAAGATGACCAAGGATATTTTGGACAGCAAGACGACATG 7930
QY 2641 MetGluValProTyrValThrProHisProTyrHisLeuGluValAlaIleLysThr 2660
DB 7931 ATGGAAATTTCCATATGTGACTCTCCCTCCCTCTTATCATTTGGAGAGCTGTCTAAAC 7990
QY 2661 TrpGluAsnThrLysGluSerProLathrHisSerSerAsnPheLeuTyrGlyThrMet 2680
DB 7991 TGGGAAAATACAAAGGAGTCTCTGCTACACATTCATCAAACTTCTGTATGGTACCATG 8050
QY 2681 ValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValleuIleCysGln 2700
DB 8051 GTTTCATACACCTGTAATCCAGATATGAACCTTCTGGGGAACCCCTGTGCTGATCTGCCAG 8110
QY 2701 GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro 2720
DB 8111 GAAGATGGAACCTTGGATGCGAGTGCACCATCTTGCAATTCATTAATGAATGTGACTTGCT 8170
QY 2721 ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetClySerAlaValGln 2740
DB 8171 ACTGCTCTCGAAAATGGCTTTTGGCTTTCACAGAGACTAGCATGGGAAGTGTGTGCAG 8230
QY 2741 TyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn 2760
DB 8231 TATAGCTGTAAACCTGGACACATTCCTTGTGGGTCTGACTTAAGGCTTTGTCTAGAGAA 8290
QY 2761 ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro 2780
DB 8291 AGAAAGTGGAGTGGTCTCCCGCTGTCAGGCTGTGAAGCCATTTTCATGCAAAAAGCCAAATCCA 8350
QY 2781 ValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGlu 2800
DB 8351 GTCATGAATGGATCCATCAAGGAAGCAACTACATACCTGAGCAGCTGTGTACTATGAG 8410
QY 2801 CysAspProGlyTyrValLeuAsnGlyThrGluAtqThrCysGlnAspAspLysAsn 2820
DB 8411 TGTGACCCCGGATATGTGCTGAATGGCTAGAGGAGAACATGCCAGGATGACAAAAAC 8470
QY 2821 TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla 2840
DB 8471 TGGATGAGGATGAGCCATTTGCAATCTCTGTGACTGTCAGTTTCACTCCAGTCTCAGCT 8530
QY 2841 AsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsn 2860
DB 8531 AATGGCCAGGTGAGAGGAGACGAGTACATTTCCAAAAAGAGATTGAATACACTTGGCAAT 8590
QY 2861 GluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSer 2880

DB 8591 GAAGGGTTCTTGTCTTGGAGGAGCCAGGAGTGGGTTTGTCTTGCCCAATGGAAGTTGGAGT 8650
QY 2881 GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly 2900
DB 8651 GGAGCCACTCCCGACTGTGTGCTGTGCAGATGTGCCACCCGCCCAACAACCTGGCCAAATGG 8710
QY 2901 ValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGly 2920
DB 8711 GTGACGGAAGGCTGGACTATGGCTTATGAAAGGAAGTAACATTCCTCATGTCTATGAGGGC 8770
QY 2921 TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu 2940
DB 8771 TACATCTTTCACGGGTCTCCAAAACCTCACCTGTCACTAGTCAAGTGGCAACTGGGATGACAG 8930
QY 2941 IleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro 2960
DB 8831 ATTTCTCTCTGTAAACCACTCAACTGTGGACTCTCTGAAGATCTTGCCCATGGTTTCCCT 8990
QY 2961 AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyrLys 2980
DB 8891 AATGGTTTTTCTTATTCATGCGGGGCCATATACAGTATACAGTGTCTTCTTGGTTATAAG 8950
QY 2981 LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerPro 3000
DB 8951 CTCCATGGAAATTCATCAGAAGGTGCTCTCCAATGGCTCTCTGGAGTGGCAGCTCACCT 9010
QY 3001 SerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThr 3020
DB 9011 TCTCTGCTGCTTGCAGATGTTCCACCCAGTAATTAATATGAATCTGTCAATGGGACA 9070
QY 3021 AspPheAspCysGlyLysAlaIleGlnCysPheLysGlyPheLysLeuLeuGly 3040
DB 9071 GATTTTGACTGTGGAAGGACAGCCCGATTCAGTGTCTTCAAAAGGCTTCAAGCTCTCAGGA 9130
QY 3041 LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu 3060
DB 9131 CTTTCTGAATCCTCTGAGACCCAGATGCCAGTGGAGCTCTGGGTCCCCCACTGTGAA 9190
QY 3061 HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer 3080
DB 9191 CACACTTCTTGTGGTTCTCTTCCAAATGATACCAATCGCTTCATCAGTGCAGACAGCTCT 9250
QY 3081 TrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSer 3100
DB 9251 TGGAAAGGAAAATGTGATTAACCTTACAGCTGCAGGTCTGGATATGTCTACAAAGGACAGTTCA 9310
QY 3101 AspLeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeu 3120
DB 9311 GATCTGATTTGTACAGAAAGGGGTATGGAGCCAGCTTATCCAGCTCTGTGAGCCCTTG 9370
QY 3121 SerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyr 3140
DB 9371 TCTGTGGGTGCCACCGCTCTGTGCCAATGCAGTGGCAACTGGAGAGGCACACACTAT 9430
QY 3141 GluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThr 3160
DB 9431 GAAAGTGAAGTGAACCTCAGATGTCTGGAAGGTATATCCATGATGATACAGATACATA 9490
QY 3161 PheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLys 3180
DB 9491 TTTCACTGTACAGAAAGTGTGCTGCTGCTGCTGCTGAGAGAAATCTCTGAGTCTCTAAAAA 9550
QY 3181 CysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsn 3200
DB 9551 TGTCTCTCTCCCGAAAAACATAACACATATATCTTGTTCATGGGACAGATTTTCACTGTGAAT 9610
QY 3201 ArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerVal 3220
DB 9611 AGGCAAGTTTCTGTGTCTATGTGCAAGGATATACCTTTTGGAGGAGTTTAACTATCAGTA 9670
QY 3221 CysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSer 3240

Db 9671 TGTCAGCTTGATGGAACCTGGAGACCCACCAATCTCCGATGAATCTTGCACTCCAGCTTTCT 9730
Qy CysGlyValProGluSerProGluHisGlyPheValValGlySerIysThrPheGlu 3260
Db 9731 TGTGGAAACCTGAAAGTCCAGAACATGGATTTGTGTGGCAGTAATAACACCTTTGAA 9790
Qy SerThrIleIleValGlnCysGluProGlyValGluLeuGluGlyAenArgGluArgVal 3280
Db 9791 AGCACAATATTATCAGTGTGAGCTGGCTATGACTAGAGGGGACACAGGAAACGGCTC 9850
Qy CysGlnGluAenArgGlnTrpSerGlyValValAlaIleCysIysGluThrArgCysGlu 3300
Db 9851 TGGCAGGAGAACAGACAGTGTGAGTGGAGGCTGCAATATGCAAGAGACAGCTGTGAA 9910
Qy ThrProLeuGluPheLeuAenGlyValAlaAspIleGluAenArgThrThrGlyProAen 3320
Db 9911 ACTCCACTTGAATTTCTCAATGGGAAGCTGACATTTGAAGAACAGGACGCTGACCAAC 9970
Qy ValValIleValSerCysAenArgGlyValSerLeuGluGlyProSerGluAlaHisCysThr 3340
Db 9971 GTGGTATATCTTGCACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCACACTGCACA 10030
Qy GluAenGlyThrTrpSerHisProValProLeuCysIysProAenProCysProValPro 3360
Db 10031 GAAATGGAACTGGAGCCACCCAGTCCCTCTCTGCAACCAATCCATGCCCTGTCTCT 10090
Qy PheValIleProGluAenAlaLeuLeuSerGluIysGluPheValAlaAspGlnAenVal 3380
Db 10091 TTGTGATTTCCCGAGAAATGCTCTGCTGTCTGAAGAGAGTGTATGTATGATCATCAATGTG 10150
Qy SerIleIysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAenPro 3400
Db 10151 TCCATCAATGTHAGGAGTGTCTGCTGACAGGCGCCATCATTAATGATGCAACCC 10210
Qy AspGluThrTrpThrGlnThrSerAlaLysCysGluIysIleSerCysGlyProProAla 3420
Db 10211 GACGAGCTGGACACAGACAGCGCCAAATGTGAAGAAATCTCATGTGTCCACCACT 10270
Qy HisValGluAenAlaIleAlaArgGlyValHisGlyGlnThrGlyAspMetIleThrThr 3440
Db 10271 CAGCTAGAAATGCAATTTGCTGAGGCGTACATTAATATGAGACATGATCACCCTAC 10330
Qy SerCysThrSerGlyThrMetLeuGluGlyPheLeuArgSerValCysLeuGluAenGly 3460
Db 10331 TCATGTACAGTGCATACATGTTGAGGGTCTTCTGAGGAGTGTGTTGTTAGAAATGGA 10390
Qy ThrTrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAenGlyGly 3480
Db 10391 ACATGACATCACCTCTCTATTGTCAGAGCTGTCTGTGATTTCCATGTCAGAAATGGGGC 10450
Qy IleCysGlnArgProAenAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGlu 3500
Db 10451 ATCTGCCAACGCCCAATGCTTGTCTGTCAGAGGGCTGGAAATGGGGGCCCTCTGTGAA 10510
Qy GluProIleCysIleLeuProCysLeuAenGlyValArgCysValAlaProIleThrGlnCys 3520
Db 10511 GAACCAATCTGCATTTCTCTCTGTCGACGAGGTGCTGTGTGGCCCTTACCAGTGT 10570
Qy AspCysProProGlyThrThrGlySerArgCysHisThrAlaValCysGlnSerProCys 3540
Db 10571 GACTGCCCGCTGGCTGGACGGGTCTGCTGTCTATACAGCTGTGTGTGCTCCAGTCTCCCTGC 10630
Qy LeuAenGlyGlyIysCysValArgProAenArgCysHisCysLeuSerSerThrThrGly 3560
Db 10631 TTAATATGGTGGAAATATGTTAAGACCAACCAACCGATGTCACGTCTTTCTTCTTGACGGGA 10690
Qy HisAenCysSerArgIysArgArgThrGlyPhe 3571
Db 10691 CATACCTGTTCCAGGAAAGAGGACTGGGTTT 10723

RESULT 2
AX375161
LOCUS

DEFINITION Sequence 1 from Patent WO0210388.
ACCESSION AX375161
VERSION AX375161.1 GI:19169916
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Welcher, A.A. and Elliot, G.S.
TITLE C3b/4b complement receptor-like molecules and uses thereof
JOURNAL Patent: WO 0210388-A 1 07-FEB-2002;
Angen Inc. (US)
FEATURES
source 1..10878
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 0 Length: 10878
Pred. No.: 0
Score: 19958.00 Matches: 3566
Percent Similarity: 99.94% Conservative: 3
Best Local Similarity: 99.86% Mismatches: 2
Query Match: 99.92% Indels: 0
DB: 6 Gaps: 0
US-09-977-053-4 (1-3571) x AX375161 (1-10878)
Qy 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
Db 11 ATGTGGCTCGCTGGCTTTTGTGTGGGTCTGGGGTCTGGGGTGGGGGACC 70
Qy 21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
Db 71 TTTTCAGCAGATGTCCTCGCTGGCAATTTTCAGCTTCCGCTTCCCGAGACCGCGGCC 130
Qy 41 GlyValProGlySerIleProAlaProAlaProGlyValAspGluAlaAlaGlySerArg 60
Db 131 GGGGCCCCGGGAGTATCCCCCGCCGCCCTCTGGCAGCAGCGGGGGAGCAGA 190
Qy 61 ValGluArgLeuGlyGlnAlaPheArgArgArgValArgLeuLeuArgGluLeuSerGlu 80
Db 191 GTGAGGGCTGGCCAGCGGCTTCGGCGACGCTGCGGTCTGCGGAGCTCAGCGAG 250
Qy 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg 100
Db 251 CGCTTGGAGCTTGTCTTCTGTGTGATGATTCGTCCAGCGTGGCGGAAGTCAACTTCGCG 310
Qy 101 SerGluLeuMetPheValArgLeuLeuSerAspPheProValValProThrAlaThr 120
Db 311 AGCGAGCTCATGTTCGTCCGCAAGCTGCTGTCCGACTTCCCGGTGGTGGCCAGGCCACG 370
Qy 121 ArgValAlaIleValThrPheSerSerLysAenThrValValProArgValAspThrIle 140
Db 371 CGGTGGCCATCTGACTTCTCTGTCCAGACACTACGTGTGGTGGCGCGCTCATTTACATC 430
Qy 141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIle 160
Db 431 TCCACCGCGCGCGCGCCAGCACAAGTGGCGCTGCTCTCTCCAAAGAGATCCCTGCCCATC 490
Qy 161 SerThrArgGlyGlyGlyThrThrLysGlyValaPheGlnGlnAlaAlaGlnIleLeu 180
Db 491 TCTTACCGAGTGGCGGCACTACACCGAGGCGCTTCCAGCAAGCGCGCAATTTCTT 550
Qy 181 LeuHisAlaArgGluAenSerThrLysValValPheLeuIleThrAspGlyThrSerAen 200
Db 551 CTTTCATGCTAGAGAAACTCAACAAAGTTGTATTTCTCATCACTGATGGATATTTCCAA 610
Qy 201 GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThr 220
Db 611 GGGGGAGACCTTAGACCAATTTGACGGCTCTCTGCGAGATTCAGGATTCAGGATCTTCACT 670

QY 221 PheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
DB 671 TTGTGGCATATGCAAGGGAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 730
QY 241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaIleAlaArgAlaLeu 260
DB 731 GAGCACTGTACTGCTACACAGTTTTGAAGAAATTTGAGGCTTTAGCTCGCGGGCATTTG 790
QY 261 HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu 280
DB 791 CATGAGATCTACCTTCTGGAGTTTATTCAGATGATATGTTCCACTGCTTATCTT 850
QY 281 CysAspGluGlyIleAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
DB 851 TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGGAAGCTGCAAAATGTTGGGACACACACA 910
QY 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyValGlyLeuGlnTyrGluCys 320
DB 911 GGCCATTTTGAAGTGCATCTGTGAAGAGGGTATTTACGGGAAGGNTTGCAGATATGAATGC 970
QY 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyIleSerSerCys 340
DB 971 ACAGCTTGGCCATCGGGACATACAACTGAAGGCTCACAGGAGGATCAGCAGTTGC 1030
QY 341 IleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCys 360
DB 1031 ATTCCATGCTCTGATGAAATCACACTCTCCACCTGGGAAGCACATCCCTGAAGACTGT 1090
QY 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
DB 1091 GTCTGGAGAGGGATACAGGGGATCTGGCCAGACCTGTGAACCTTGTCCACTGCCCTGCC 1150
QY 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
DB 1151 CTGAAGCTTCCGAAAATGTTACTTTATCCAAACACTTGCACACACCTTCAATGCA 1210
QY 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
DB 1211 GCCTGTGGGGTCCGATGTCACTGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1270
QY 421 LeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
DB 1271 CTACCAATGGTTTGTGGTCCGGTTCAGAGAGCTACTGCAGAGTAAGAACAATGCTCTCAT 1330
QY 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1331 CTCCGCCAGCCGAACATGCGCCACATCAGCTGTCTTACAAAGGMAATGTTATATAAGACA 1390
QY 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1391 ACATGTTTGGTTCCTGTGATGAAGGGTACAGACTAGAAAGGCGATGAATAGCTTACTGT 1450
QY 481 GlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1451 CAAGAAACAGCCAGTGGATGGCCGAACCCCGGTGTGTGAGCGCCACATGTTCCACC 1510
QY 501 PheGlnMetProLysAspValIleSerProHisAsnCysGlyLysGlnProLys 520
DB 1511 TTTCAGATGCCCAAGATATCAT 1570
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1571 TTGTGGAGCGATCTGCTATGTAAGTTGTCCCGCAAGGGTTTCAATTTATCTGGAGTCAAGAA 1630
QY 541 MetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAlaValCysLys 560
DB 1631 ATGCTGAGATGTACCACTTCTCGAAATATGAAATGTCGAGTTTCAGGAGCTGTGTGAAA 1690
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
DB 1691 GACGTGGAGGCTCCTCAATCAACTGCTCTAAGGACATAGAGGCTTAAGACTCTGGAAACAG 1750

QY 581 GlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 1751 CAAGATTTCTGCCAATGTTTACCTGGCAGATTCCAAACGCTAAAGACAACCTCTGTGTGAAGA 1810
QY 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
DB 1811 GTGTGATCCAGTTTCATCCAGCTTTCACCCACCTTACCTTTCCAGTTGGAGATGTT 1870
QY 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
DB 1871 GCTATCGTATACACGGCAACTGACCTATCCGGCAACAGCGCCAGCTGCAATTTTCCATATC 1930
QY 641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
DB 1931 AAGGTTATTGATGTCAGAACCTCTGTCATAGACTGGTGCAGATCTCCACCTCCGTCAG 1990
QY 661 ValSerGluLysValHisAlaIleSerTyrAspGluProGlnPheSerAspAsnSerGly 680
DB 1991 GTCTGGAGAGGTACATGCCGACAGCTGGATGAGCTTCAGTTCTCAGACACACTCAGGG 2050
QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
DB 2051 GCTGAATTTGTTTACCTACCAAGAGTTCATACCAAGGAGACCTTTTCCCTCAAGGGAGACT 2110
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
DB 2111 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATTAACAGGACATGTGATATCCATATT 2170
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
DB 2171 GTCATAAAGGTTCTCCCTGTGAAATTCATTCACACTGTAAATGGGGATTTTATATGC 2230
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
DB 2231 ACTCCAGATAATACCTGGAGTCAACTGTACATAAATTCCTTGTGGGGGCTATGATTTTCA 2290
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
DB 2291 GAAGGCTTACTGACAAAGTATTTATGCTTATGAAGTGGCGCTGGAACCAACACATAT 2350
QY 781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
DB 2351 ACCACTGAATGCCAGACTGTGCCAAAAAAGCTTTTGTCTAACCAACCGGTTCAAGTCTTT 2410
QY 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
DB 2411 GAGATGTTCTACAAAGCAGCTGCTGTGATGACACACAGATCTGATGAAGAAGTTTCTGAA 2470
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
DB 2471 GCATTTGAGACGACCTCGGAAATAATGTTCCCATCATTTTGTAGTATGCAAGGAGACTT 2530
QY 841 AspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
DB 2531 GACTGCAGACTGGAGAGAACCTGCACAAAAAATAATTTGCTAGATAATAATTTATGACTAT 2590
QY 861 GluAsnGlyPheAlaIleGlyProGlyTyrGlyValAlaAlaAsnArgLeuAspTyrSer 880
DB 2591 GAAATTTGGCTTTGCAATTGACACAGGTGGTGGGTGGAGCTAAATAGGCTGGATTAATCTCT 2650
QY 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
DB 2651 TAGATGATCTCTGACACACTGTGCAGAAACAGCCACAGCATCCGCAATGCCAAGTCC 2710
QY 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
DB 2711 TCACGGATTTAAAGAAAGTGGTCCCATTTCTGACTATAAAATTAAGTTAATTTTAAACATC 2770
QY 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnGln 940
DB 2771 ACAGTAGTGTGCCATTTACCGATGAAGAAATGATACCTTGAATGGGAAATTCAGCAA 2830
QY 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960

1681 ProLeuProHisCysGluAraIleSerCysGlyValProProProLeuGluAsnGlyPhe 1700
1701 HisSerAlaAspPhePheThrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
5051 CCACTTCTCCTGTAACGCAATAGCTGTGGGTGCGCACCTCTTTGGAGAAATGGCTTC 5110
1701 HisSerAlaAspPhePheThrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyr 1720
5111 CATTACCGCATGACTTCTATGCTGGCAGCACAGTAACCTACCAGTGCACAAATGGCTAC 5170
1721 TyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSer 1740
5171 TATCTATTGGTGACTCAGAGATGTTCTGTACAGATAATGGAGCTGGAAAGCGGCTTCA 5230
1741 ProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSer 1760
5231 CCATCTGCTTCATGTCGATGAGTGTGCGAGTGGATCAGATTGTAGTACGATGCTTCT 5290
1761 CysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGly 1780
5291 TGCCTGAACGTAGATGGATCTTACATATGTTCTATGTGTCCACCGTACACAGAGATGGG 5350
1781 LysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHisSer 1800
5351 AAAAATGTGCAGAACCTATAAATGTAGGCTCCAGGAATCCGGAATATGGCAATGGCCACTCC 5410
1801 SerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGln 1820
5411 TCAGGTGAGATTTATACAGTAGGTGCGCAAGTCACATTTCTGTCAGGAAGGATACCA 5470
1821 LeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisLeuIlePro 1840
5471 TTGATGGAGTAAACCAAAATCACATGTTTGGAGTCTGGGAATGGAATCATCTAATACCA 5530
1841 TyrCysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGlu 1860
5531 TATTGTAACTGTTTCATGTGTAAACCGCTATTCAGAAATGTTTGCATTGAGGAG 5590
1861 LeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAla 1880
5591 TTAGCATTTACTTTTGGCAGCAAGTGACATATAGGTGTATAAATGAAGATATACTCTGGCC 5650
1881 GlyAspLysGluSerSerCysLeuAsnSerSerTrpSerHisSerProProValCys 1900
5651 GGTGATAAAGAAATCATCTGCTGTCTGTCTTGGAGTCAATCCCTCTCTGTGTGT 5710
1901 GluProValLysCysSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGly 1920
5711 GAACCACTGAGTGTCTAGTCCGGAATATATAAATGAAGATATATTTTGGATGGG 5770
1921 LeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyPro 1940
5771 CTTACCTTACCTTTCTACTGTCATCATATTCATGCGATACAGGATACAGCTTACAGGGCCCT 5830
1941 SerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHisLeu 1960
5831 TCCATATTGAATGACCGGCTTCGGCATCTGGGACAGAGCGCCACTGCTGTACCTC 5890
1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
5891 GTCTTCTGTGGAGNACCACCTGCCATCAAGATCTGTCTATTACGGGAATAAATCACT 5950
1981 PheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThr 2000
5951 TTCAGGAACACCGTCACTTACATTCGAAAGAAGGCTATATCTTGTGCTGTGTGACACC 6010
2001 IleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSer 2020
6011 ATTGAATGCTTGGCGGCGGCAAGTGGAGTAGAAGTGCACAGAGTGCCTGCTGTCTCC 6070
2021 CysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGly 2040
6071 TGTGATGAGCCACCATTTGTGGACCAACGCTCTCCAGAGAGCTGCCCATCGCTCTTTGGA 6130

2041 AspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeu 2060
6131 GACATTTGCTTACTACTCTCTCTGATGGTTACAGCCTAGCAGCAATATCCAGCTTCTC 6190
2061 CysAsnAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIleAla 2080
6191 TGCATGCCCHGGCAGGTGGGTACCCCAAGAGTCAAGACATGCCCTGTTGTATAGCT 6250
2081 HisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
6251 CATTTCTGTGAAAAAOCCTCCATCGGTTTCTATAGCATCTTGGAAATCTGTGAGCAAGCA 6310
2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
6311 AAAATTTGCAGCTGGCTCAGTTGTGAGCTTTAAATGCATGGAGGCTTTGTACTGAACACC 6370
2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
6371 TCAGCAAGATTTGAATGTATGAGAGGTGGCGAGTGGAAACCTTCCCCCATGTCCATCCAG 6430
2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
6431 TGCATCCTGTGCGGTGTGGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGATCA 6490
2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
6491 AACTACAGTTTGGAGCCATGTTGGCTTACAGCTGCACAGGGGCTTCTACATCAAGGG 6550
2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis 2200
6551 GAAAGAGAGAGCACCTGCGAAGCCACAGGCGAGTGGAGTAGTCTCTATACCGAGGTGCCAC 6610
2201 ProValSerCysGlyGluProProLysValIleAsnGlyPheLeuGluHisThrThrGly 2220
6611 CCGGTATCTTGTGTGAAACCACTTAAGTTCAGAAATGGCTTCTTGGAGCATACACTGGC 6670
2221 ArgIlePheGluSerGluValAlaTyrGlyCysAsnProGlyTyrLysSerValGlySer 2240
6671 AGGATCTTTGAGAGTGAAGTGAGGTATCAGTGAACCCGGGCTATAAGTCAAGTGGAGT 6730
2241 ProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVal 2260
6731 CCTGTATTTGTCTGCCAAGCCATCGCCACTGGCACAGTGAATCCCTCTCATGTGTGT 6790
2261 ProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPhe 2280
6791 CCTCTCGACTGTGGAAAACTCCGCCATCCAGATCCAGATGGCTTTCATGAAGAGAAACTTT 6850
2281 GluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSer 2300
6851 GAAGTAGGGTCCAGGGTTCAGTTTCTGTATGAGGGTTATGAGCTTGTGTGGACAGT 6910
2301 SerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPro 2320
6911 TCTTGGACATGTCCAGAAATCTGGCAATGGAATGAAGTCAATCCAAAGTGCATGCT 6970
2321 AlaLysCysProGluProProLeuLeuGlnAsnGlnLeuValLeuLysGluLeuThrThr 2340
6971 GCCAAGTGGCCGAGCGCCCTCTTGGAAACACAGCTAGTAGTATTAAAGAGATTGACACC 7030
2341 GluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerVal 2360
7031 GAGTAGAGATTGTGACATTTCTGTATAGAGAGGCGATGTCTCTGCAAGGCCCTCTCTC 7090
2361 LeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeu 2380
7091 CTGAATGTCTGCCATCCAGCAATGGAATGACTCTTTCCTCTGTGTGTGAAGATTGTCTT 7150
2381 CysThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPhe 2400
7151 TGTACCCCACTCCCTCAATTTCTTGGTGTGCCCATTCCTTCTCTGCTCTTCTTCTT 7210
2401 GlySerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThr 2420

Db	7211	GGAGTACTGTCAAGTATTTCTGTGTAGTGGGTTTTTCTTAAGAGAAATTTCTACCA	8350	AGAAAGTGGAGTGGCTCCCTCCACGCTGTGAAGCCATTTTCATGCAAAAAGCCAAATCCA
Qy	2421	LeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCys	2800	ValMetAsnGlySerIleLysGlySerAsnTrpThrTrpLeuSerThrLeuTrpGlu
Db	7271	CTCTGCCAACCTGATGGCACCTGGAGCTCTCCACATGCCAGAAATGTGTTCAGATGAATGT	8410	GTCAATGAATGGATCCATCAAGAGAACCACTACACATACCTTGAGCAGCTGTGTACTATGAG
Qy	2441	ProGlnProGluGluIleProAsnGlyIleLeuAspValGlnGlyLeuAlaTrpLeuSer	2820	CysAspProGlyTrpValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsn
Db	7331	CCCCAACCTGAGGAAATCCCAATGGAAATCAATGTGTGAGGCCCTTGCCCTATCTCAGC	8470	TGTGACCCCGGATATGTGTGAATGGCACTGAGAGAGAAACATGTCAGGATGACAAAC
Qy	2461	ThrAlaLeuTrpThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCys	2840	TrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSerAla
Db	7391	ACAGCTCTCTATACCTGCAAGCCAGGCTTGAATTTGGTGGGAATACATCCACCCCTTGT	8530	TGGATGAGGATGAGCCCATTTGCATTTCTGTGGACTGCAGTTTCAACCCCACTCTCAGCC
Qy	2481	GlyGluAsnGlyHisTrpLeuGlyGlyIleProThrCysLysAlaIleGluCysLeuLys	2860	AsnGlyGlnValArgGlyAspGluTrpThrPheGlnLysGluIleGluTrpThrCysAsn
Db	7451	GGAGAAATGGTCACTGGCTTGGAGGAAACCAACATGTAAGCCATTGAGTGCCTGAAA	8590	AATGGCCAGGTGAGAGAGAGAGTACATTTCCAAAAGAGATTTGAATACACTTGGCAAT
Qy	2501	ProLysGluIleLeuAsnGlyIlePheSerTrpThrAspLeuHisTrpGlyGlnThrVal	2880	GluGlyPheLeuLeuGluGlyValArgSerArgValCysLeuAlaAsnGlySerTrpSer
Db	7511	CCCAAGAGATTTTGAATGGCAAAATTTCTTTACAGGACCTTACACTATGGACAGCCGTT	8650	GAAGGTTCTTGTCTTGAGGAGCCAGAGTCCGGTTTGTCTTGCCAATGGAAATGGAGT
Qy	2521	ThrTrpSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGlu	2900	GlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGly
Db	7571	ACCTACTCTTGCACCGAGGCTTTCGTCGAAGTCCAGTGCCTTGACCTGTTAGAG	8710	GGAGCCACTCCCGACTGTGTGCTGTGATGTGCACCCCGCCACAACTGGCCAAATGGG
Qy	2541	ThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGln	2940	TyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGlu
Db	7631	ACAGGTGATTTGGGATGTAGATGCCCATCTTGCATATGCCATCCACTGTGATTTCCCA	8770	GTGACGGAAGGCCCTGGACTATGGCTTCATGAAGAGAGTAAATTCCTCATGTCTGAGG
Qy	2561	ProIleGluAsnGlyPheValGluGlyValAspTrpSerTrpGlyAlaIleIleTrp	2980	ValThrGluGlyLeuAspTrpGlyPheMetLysGluValThrPheHisCysHisGluGly
Db	7691	CCCATTTGAATGGTTTTTGTAGAGGGTGCAGATTACAGCTATGTGTGCCATATCTTAC	8830	TACATCTTGACGGTGTCTCCAAACTCACCTGTGATGAGTGGCACTGGGATGCAGAG
Qy	2581	SerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGly	2960	IleProLysCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhePro
Db	7751	AGTTGCTCTCCCTGGGTTTCAGGTGGCTGTCTATGCCATGCAGACTGTGAAGAGTCAG	8890	ATTCCTCTCTGTAAACCACTGAGTGGACCTCTCTGAAGATCTTGGCCATGGTTTCCCT
Qy	2601	TrpSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAsp	2980	AsnGlyPheSerPheIleHisGlyGlyHisIleGlnTrpGlnCysPheProGlyTrpLys
Db	7811	TGCTCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCTCTCCCTCATATAGT	8950	AATGGTTTTCTTTATTCATGGGGCCATATACAGTATCAGTGTCTTCTGGTTTATAG
Qy	2621	PheGlyAspCysThrLysLeuLysAspGlnGlyTrpPheGluGlnGluAspMet	3000	LeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerPro
Db	7871	TTTGGAGACTGTACTAACTCAAAGATGACCAAGGATATTTTGACGAAGAGACGACATG	9010	CTCCATGGAATTCATCAAGAGGTGCTCTCTCCAATGGCTCTCTGGAGTGGCAGCTCACCT
Qy	2641	MetGluValProTrpValThrProHisProProTrpHisLeuGlyAlaValAlaLysThr	3020	SerCysLeuProCysArgCysSerThrProValIleGluTrpGlyThrValAsnGlyThr
Db	7931	ATGGAGTTCATATGTGACTCTCCCTCCTCATCTTATCTTGGAGCAGTGGCTTAAAC	9070	TCCTGCTGCTTGCAGATGTTCCACACCACTGAATTTGAATATGGAACCTGTCAATGGGACA
Qy	2661	TrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTrpGlyThrMet	3040	AspPheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGly
Db	7991	TGGGAAATACAAAGGAGTCTCTGCTACACATTCATCAAACTTCTGTATGGTACCATG	9130	GATTTTGACTGTGGAAGGCGCCGGATTCAGTGTCTTCAAGGGCTTCAGCTCTCAGGA
Qy	2681	ValSerTrpThrCysAsnProGlyTrpGluLeuLeuGlyAsnProValLeuIleCysGln	3060	LeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGlu
Db	8051	GTTTTCATACACCTGTAAATCCAGGATATGAATCTCTGGGGAACCTCTGCTGTGATC	9190	CTTTCTGAATTCATCTGTGAAGCCGATGGCCAGTGGAGCTCTGGGTTCCTCCCACTGTGAA
Qy	2701	GluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuPro	3080	HisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSer
Db	8111	GAAGATGGAATTTGAATGGCAGTGCACCATCTCTGCATTTCAATTTGAATGTGACTTGCCT	9250	CACACTTCTTGTGTCTCTTCCAAATGATACCAATTCGTTTCATCAGTGAGACCACTCT
Qy	2721	ThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGln	3100	TrpLysGluAsnValIleThrTrpSerCysArgSerGlyTrpValIleGlnGlySerSer
Db	8171	ACTGCTCTCTGAAATGGCTTTTTCGTTTTTACAGAGACTAGCATGGGAAGTGTGTGAG	9310	TGGAGGAAATGTGATAACTTACAGCTGCAGGTCTGGATATGTCTATCAAGGCACTTCA
Qy	2741	TrpSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsn	3120	AspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpValCysGluProLeu
Db	8231	TATAGCTGTAAACCTGGACACATTTCTAGTGGGCTCTGACTTTAGGCTTTGTCTAGAGAT	9370	GATCTGATTTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCAGTCTGTGAGGCCCTTG
Qy	2761	ArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnPro	9430	TCCTGTGGGTTCCCACTGTCTGTCGCAATGTCAGTGGCACTGGAGAGGCAACACACTAT

[illegible]

Qy	1564	PheSerProLaGluSerPheValGlySerIleSerGlnLeuAsnLeuTrpAspTyrVal	1586
Db	4757	TTCAACCCGGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTCTGGGACTATGTC	4816
Qy	1584	LeuSerProGlnGlnValIysSerLeuAlaThrSerCysProGluGluLeuSerIysGly	1603
Db	4817	CTGTCTCCACAGCAGGAGGAAAGTCACTGGCTACCTCTCCGCCAGAGGAACTCAGTAAGGA	4876
Qy	1604	AsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyIysValIysIleAspSer	1623
Db	4877	AACGTGTAGTAGGCCCTGATTTCTTGTTCAGAAATGTGGGGAAGTGAAGATCGATTCT	4936
Qy	1624	LysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeuArgThr	1643
Db	4937	AAGAGCATATTTTGTCTGATTTGCCACGCTTGGAGGGTCAGTGCCTCATCTGAGAACT	4996
Qy	1644	AlaSerGluAspLeuIysProGlySerIysValAsnLeuPheCysAspProGlyPheGln	1663
Db	4997	GCATCTGAAGAATTTAAACACGAGTTCCAAAGTCAATCTGTTCTGTGAACAGAGCTCCAG	5056
Qy	1664	LeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnProLeuPro	1683
Db	5057	CTGGTCGGGAACCCCTGTGCAGTACTGTCTGATTCAGACAGTGGGACACCACTCCCC	5116
Qy	1684	HisCysGluArgIleSerCysGlyValProProLeuGluAsnGlyPheHisSerAla	1703
Db	5117	CACGTGAAACGCAATTGCTGTGGGGTGCCACCTCTCTTTGGAGAATGCTTCCATTACGCC	5176
Qy	1704	AspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeu	1723
Db	5177	GATGACTTCTATGCTGGCAGCACAGTAAACCTTACCAGTGCACCAATGGCTACTATCTATTG	5236
Qy	1724	GlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCys	1743
Db	5237	GGTGACTCAGAGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCTTTCACCATCTCTGC	5296
Qy	1744	LeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsn	1763
Db	5297	TTAGATGTGATGATGATGTCAGTGTGGATCAGATTGTAGTGAGCATGCTTCTTGCTCGAAC	5356
Qy	1764	ValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyIysAsnCys	1783
Db	5357	GTAGATCGATCTTACATATGTTGATGTGTCCACCGTACAGAGAGATGGGAAAACTGT	5416
Qy	1784	AlaGluProIleIysCysLysAlaProGlyAsnProGluAsnGlyHisSerSerGlyGlu	1803
Db	5417	GCAGAACTATTAATGTTAGGCTCCAGGAATATCCGGAAATGGCCACTCTCTAGGTGAG	5476
Qy	1804	IleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGly	1823
Db	5477	ATTATATACAGTAGGTGCGGAAGTCACATTTTGTGTTCAGAGAGATACCAAGTTGATGGGA	5536
Qy	1824	ValThrIysIleThrCysLeuGluSerGlyGluTrpAsnHisIleuIleProTyrCysLys	1843
Db	5537	GTAAACCAAAATCACATGTTTGGAGCTTCGGAGAATGGAATCATCTAATACCATAATGTAAA	5596
Qy	1844	AlaValSerCysGlyIysProAlaIleProGluAsnGlyCysIleGluGluLeuAlaPhe	1863
Db	5597	GCGTGTTTCATGTGTAAACCGCTATTTCAGAAAATGGTTGCATTGAGGAGTTAGCAATT	5656
Qy	1864	ThrPheGlySerIysValThrTyrArgCysAsnIysGlyTyrThrLeuAlaGlyAspLys	1883
Db	5657	ACTTTTGGCAGCAAAAGTGACATATAGGTGTAATAAAGGATATATCTCTGGCCGGTGATAAA	5716
Qy	1884	GluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProProValCysGluProVal	1903
Db	5717	GAATCATCTGCTGTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5776
Qy	1904	LysCysSerSerProGluAsnIleAsnAsnGlyIysTyrIleLeuSerGlyLeuThrTyr	1923
Db	5777	AAGTGTTCTAGTCCGGAATAATATAATAATGAATAATATATTTTGTAGTGGGCTTACTCTAC	5836
Qy	1924	LeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIle	1943

Db	5837	CTTTCTACTGCATCATATTCTCATGCCGATACAGGATACAGCTTTACAGGCGCCTTCACATTATT	5896
Qy	1944	GluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHisLeuValPheCys	1963
Db	5897	GAATGCACGGCTTCTGGCATCTGGGACAGAGCGCACCTGCCTGTCACTCGTCTTCTGT	5956
Qy	1964	GlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsn	1983
Db	5957	GGAGAACCACTGCGCATCAAGAATGCTGTCAATTACGGGGAATAAATCTCACTTTCAGGAAC	6016
Qy	1984	ThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCys	2003
Db	6017	ACCGTCACTTACACATTGCAAGAAGGCTATCTCTTGCTGCTGTGACACCATGTGAATGC	6076
Qy	2004	LeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaValSerCysAspGlu	2023
Db	6077	CTGGCCAGCGCAAGTGGAGTAGAGTGACACGAGTGCTGCTGCTCTCTGTGATGAG	6136
Qy	2024	ProProfileValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAspIleAla	2043
Db	6137	CCACCCATTGTGACCACGCGCTCTCCAGAGACTGCCCATCGGCTCTTTGGAGACATTGCA	6196
Qy	2044	PheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAla	2063
Db	6197	TTCTACTACTGCTCTGTATGGTTACAGCTTAGCAGACAAATTCACGAGCTTCTCTGCAATGCC	6256
Qy	2064	GlnGlyLysTyrValProGluGlyGlnAspMetProArgCysValIleAlaHisPheCys	2083
Db	6257	CAGGCGANGTGGTACCCCCAGAGAGGTCAMAGCATGCCCGCTGTGTATGCTCATTTCTGT	6316
Qy	2084	GluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAlaLysPheAla	2103
Db	6317	GAATAACCTCCATCGGTTTCTCTATAGCATCTTGGAAATCTGTGAGCAAGCAAAATTTGCA	6376
Qy	2104	AlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLys	2123
Db	6377	GCTGGCTCAGTTGTGAGCTTTAAATGCATGGAAGGCTTTGTACTGGAACACCTCAGCAAG	6436
Qy	2124	IleGluCysMetArgGlyGlyGlnTyrAsnProSerProMetSerIleGlnCysIlePro	2143
Db	6437	ATTGAAATGATAGAGGTGGCGAGTGGAAACCTTCCCCANGTCCATCCAGTGCATCCCT	6496
Qy	2144	ValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSer	2163
Db	6497	GTGCGGTGTGAGAGCCACCAAGCATCATGAATGGCTATGCAAGTGGATCAAACTACAGT	6556
Qy	2164	PheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyGluLysLys	2183
Db	6557	TTTGGAGCCATGTGTGGCTTACAGCTGCGAAACAGGGGTTCTACATCAAGAGGGGAAAAAG	6616
Qy	2184	SerThrCysGluAlaThrGlyGlnTyrSerSerProIleProThrCysHisProValSer	2203
Db	6617	AGCACCTGCCAAGCCACAGGGCGAGTGAGTAGTCTTATACCGACGTGCCACCCGGTATCT	6676
Qy	2204	CysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePhe	2223
Db	6677	TGTGGTGAACCACTTAAGGTTGAGATGGCTTTCTGGAGCATACAACCTGCGAGATCTTT	6736
Qy	2224	GluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPhe	2243
Db	6737	GAGAGTGAAGTAGGATTCAGTGTAAACCCGGGCTATAAGTCACTCGGAAGTCTCGTATTT	6796
Qy	2244	ValCysGlnAlaAsnArgHisTyrHisSerGluSerProLeuMetCysValProLeuAsp	2263
Db	6797	GTCGCGCAACCAATCGCCACTGGCACAGTGAAATCCCTCTGTATGTGTGTCTCTCTCGAC	6856
Qy	2264	CysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGly	2283
Db	6857	TGTGGAACCTCCCCGATCCAGATGGCTTCATGAAGAGGAGAAACTTTTGAAGTAGGG	6916
Qy	2284	SerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTyrThr	2303

Db	6917	TCCAAGGTTCAAGTTCTTTCTGTAATGAGGGTTATGAGCTTGTGTGGTGACAGTCTCTTGGACA	6977
Qy	2304	CysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysCysMetProAlaLysCys	2323
Db	6977	TGTGAGAAATCTGGCNAATGGAAATAGAAAGTCAATCCNAAGTCCATGCCCTGCCAGTGC	7036
Qy	2324	ProGluProProLeuLeuGluAAsnGlnLeuValLeuLysGluLeuThrThrGluValGly	2343
Db	7037	CCAGAGCGCGCCCTCTTGGAANAACAGCTAGTATATAAGGAGTGTGACCAACGAGGTAGGA	7096
Qy	2344	ValValThrPheSerCysLysGluGlyHisValLeuGlnGlnProSerValLeuLysCys	2363
Db	7097	GTTGTGCAATTTCTCTGTAAGAAGGGCATGCTCTGCAAGGCCCTCTGTCTGTAATATGC	7156
Qy	2364	LeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleValLeuCysThrPro	2383
Db	7157	TTGCCATCCAGCAATGGAAATGACTCTTCCCTGTTTGTAAAGATGTCTCTTGTATACCCA	7216
Qy	2384	ProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHisPheGlySerThr	2403
Db	7217	CCTCCCTTAATTTCTTTGGTGTGCCCATTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	7276
Qy	2404	ValLysTrpSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGln	2423
Db	7277	GTCAGATATCTCTGTAGTGTGGGTTTTCTTAAGAGGAAATCTACACCCCTCTGCCAA	7336
Qy	2424	ProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCysProGlnPro	2443
Db	7337	CCTGATGGCACCCTGGAGCTCTCCACTGCCAGAAATGTGTTCAGTAGAAATGTCCCACT	7396
Qy	2444	GluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTrpLeuSerThrAlaLeu	2463
Db	7397	GAGCAAAATCCCAATGGAATCATTTGATGTGCAAGGCCCTTGCTATCTCAGCAGAGCTCTC	7456
Qy	2464	TyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsn	2483
Db	7457	TATACCTGCAAGCCAGAGCTTTGAAATTTGGTGGGAAATATCTACCAACCTTTGTGGAGAAA	7516
Qy	2484	GlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLysProLysGlu	2503
Db	7517	GGTCACTGGCTTGGAGGAAACCAACATGTAAAGCCATTGATGGCTCTGAAACCAAGGAG	7576
Qy	2504	IleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSer	2523
Db	7577	ATTTTGAATGGCAAAATTTCTCTTACAGGACCTACACTATGAGCAGACCGTTACTACTCT	7636
Qy	2524	CysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAsp	2543
Db	7637	TGCAAACGAGGCTTTTGGCTTCGAAAGTCCAGAGTGCCTTGAACCTGTTTAGAGACAGGTGAT	7696
Qy	2544	TrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGlu	2563
Db	7697	TGGGATGTAGATGCCCATCTTGCAATGCCATCCACTGTGATTTCCCAACCAACCATTTGAA	7756
Qy	2564	AsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPhe	2583
Db	7757	AATGGTTTTGTAGAAGGTGCAGATTACAGCTATGTCGCCAATATCATCTACAGTTGCTTC	7816
Qy	2584	ProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSer	2603
Db	7817	CCTGGGTTTCAGGTGGCTGGTCAATGCCATGCAAGACCTGTGAGAGTCAAGTATGCTCAAGT	7876
Qy	2604	SerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPheGlyAsp	2623
Db	7877	TCCATCCCAACATGTATGCCAATAGACTGTGGCCCTCCCTCTCATATAGATTTTGGAGAC	7936
Qy	2624	CysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAspMetMetGluVal	2643
Db	7937	TGTACTAACTCAAGATGACCAGGATATTTTGAGCAAGAGACGACATGATGGAGTT	7996
Qy	2644	ProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsn	2663
Db	7997	CCATATGTGACTCTCACCCTCTTATCAATTTGGAGCTGTGCTAAACCTGGGAAAT	8056

QY	2664	ThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyr	2689
DB	8057	ACAAGGAGTCTCTCGCTACACATTCATCAAACTTTCTGTATGCTAGCATGCTTTCATAC	8116
QY	2684	ThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLeuLeuLeuLeuGluAspGly	2703
DB	8117	ACCTGTATTCGAGTATGAACCTCTCGGGAACCTGTGCTGATCTGCCAGGAAGATGGA	8176
QY	2704	ThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaPro	2723
DB	8177	ACTTGAATGACGTGCACCATCTGCATTCATTTGAATGTGATCTTGCCTACTGCTCCT	8236
QY	2724	GluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCys	2743
DB	8237	GAATAATGCTTTTTCGCTTTTACAGAGACTAGCATGGGAAGTGCTGTGCAGTATAGCTGT	8296
QY	2744	LysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrp	2763
DB	8297	AAACCTGGACACATTTCTAGCAGGCTCTGACTTAAGGCTTTGTCTAGAGAAATAGAAGTGG	8356
QY	2764	SerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnProValMetAsn	2783
DB	8357	AGTSGTGCTCTCCCAACGCTGTGAAGCCATTTTCATGCAAAAGGCCAAATCCAGTCATGAAT	8416
QY	2784	GlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspPro	2803
DB	8417	GGATCCATTCAGGAGACACTACACATCTGAGCAAGTTGTACTATGATGTGACCCC	8476
QY	2804	GlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrpAspGlu	2823
DB	8477	GGATATGTGCTGAATGGCACTGAGAGAGAGAACATGCCAGCATGCAAAAACCTGGGATGAG	8536
QY	2824	AspGluProIleCysIleProValAsnCysSerSerProProValSerAlaAsnGlyGln	2843
DB	8537	GATGAGGCCATTTTCATTTCTGTGGACTGTGAGTTCACTCCAGTCTCAGCCAAATGGCCAG	8596
QY	2844	ValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPhe	2863
DB	8597	GTGAGAGGAGACGAGTACACATTTCCAAAAGAGATTGAATACACTTGCAATGAAGGGTTC	8656
QY	2864	LeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThr	2883
DB	8657	TTGCTGTAGGGAGCCAGGAGTCGGGGTTGTCTTTGCCAATGGAATTTGGAGTGGAGCCACT	8716
QY	2884	ProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGlu	2903
DB	8717	CCCAACTGTGTGCTGTCAAGTGTCCACCCCGCCACACTGGCCCAATGGGGTATCGGAA	8776
QY	2904	GlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeu	2923
DB	8777	GGCTCTGACTATGGCTTCATGAAGGAAGTAACCATTTCCACTGTCCAGGGGCTACATCTTG	8836
QY	2924	HisGlyAlaProLysValLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeu	2943
DB	8837	CACGGTGCTCCAAACCTCACCTGTCACTAGATGGCACTGGGATGCGAGATTCCTCTC	8896
QY	2944	CysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPhe	2963
DB	8897	TGTAACCCAGTCAACTGTGGACCTCTCTGAAGATCTTTGGCCCATGGTTTCTCCCTAATGGTTT	8956
QY	2964	SerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGly	2983
DB	8957	TCCTTTTATTCATGGGGGCCATATACAGTATCACTGTCTCTCTGTTATATAGCTCCATGGA	9016
QY	2984	AsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeu	3003
DB	9017	AATTCATCAAGAGGTGCTCTCCATGGCTCTCTGGAGTGGCAGCTCACCTTCTCTGCTG	9076
QY	3004	ProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAsp	3023
DB	9077	CCTTGCAATGTTCCACACCAAGTAATGAATATGAATATGAATCTCAATGGGACAGATTTTGAC	9136

QY	3024	CysGlyLysAlaAlaargileGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGlu	3043	Db	10217	TGTAGGAGGTTTCTGCTGAGGCGCCACCGCATCATTTACCTGCAACCCCGACGAGACG	10276
Db	9137	TGTGGAAAGGCGAGCCCGGATTTCAGTGTCTCAAGAGCTTCAAGCTCTAGGACTTTCGAA	9196	QY	3404	TrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProLahHisValGlu	3423
QY	3044	IleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSer	3063	Db	10277	TGGACACAGACAGCGCCAAATGTGAAAAAATCTCATGTGGTCCACAGCTCAGTAGAA	10336
Db	9197	ATCACCTGTGAAGCGGATGGCCAGTGGAGCTCTGGGTTCCTCCCACTGTGAACACACATCTCT	9256	QY	3424	AsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyr	3443
QY	3064	CysGlySerLeuProMetIleProMetIlePheLeuSerGluThrSerSerTrpLysGlu	3083	Db	10337	ATGCAATGCTCGAGGCGTACATTTATCATATGGAGCATGATCACTTACTCATGTTAC	10396
Db	9257	TGTGTTCTCTTCCAAATGATACCAATGCTTCATCAGTGAGCAGCTCTTGGAAAGGAA	9316	QY	3444	SerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThr	3463
QY	3084	AsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIle	3103	Db	10397	AGTGATACATCTGAGGGTTCTCTGAGGAGTGTGTTGTTAGAAAAATGGAACATGSA	10456
Db	9317	ATGTCATTAATCTACAGCTGACAGTCTGGATATGTCATACAGGCGAGTTCAGATCTGATT	9376	QY	3464	SerProProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGln	3483
QY	3104	CysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSerCysGly	3123	Db	10457	TCACTCTCTATTTGAGAGCTGTCTGTGATTTCCATGTGCAGATGGGGCATCTGCCAA	10516
Db	9377	TGTACAGAGAAAGGGGTATGGAGCCAGCCCTTATCCAGTCTGTGAGCCCTTGTCTCTGG	9436	QY	3484	ArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGluGluProIle	3503
QY	3124	SerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGlu	3143	Db	10517	CGCCCAATGCTGTCTCTGTCAGAGGCTGATGGGGCGCTCTGTGAGAACCAATC	10576
Db	9437	TCCCCACCGTCTGTGCCAATGCATGCGCACTGGAGAGGCACACACTATGAAAGTGAA	9496	QY	3504	CysIleLeuProCysLeuAsnGlyGlyArgCysValAlaLysValAlaProTyrGlnCysAspCysPro	3523
QY	3144	ValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrPheThrCys	3163	Db	10577	TGCACTCTCTCTGCTGAACGAGGCTGCTGTGTGGCCCTTTACCAGTGTGACTGCCCG	10636
Db	9497	GTGAACCTCAGATGTCTGGAGGTTATACGATGGATACAGATACAGATACATTCACCTGT	9556	QY	3524	ProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGly	3543
QY	3164	GlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeu	3183	Db	10637	CTGTGCTGACGGGTCTCGCTGTATACAGCTGTTTGCAGTCTCCCTGCTTAAATGGT	10696
Db	9557	CAGAAAGATGGTGGTTCCTCGAGAGAAATCTCTGCGAGTCTCTAAAAAAATGCTCTCTC	9616	QY	3544	GlyLysCysValArgProAsnArgCysHisCysLeuSerSerTyrThrGlyHisAsnCys	3563
QY	3184	ProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnVal	3203	Db	10697	GGAAATGTGTAGAACCAACCAACGATGTCACTCTCTTCTTGTGACGGGACATACACTGT	10756
Db	9617	CCGAAACATTAACACATATATCTGTACATGGGAGCATTTTCAGTGTGAAATAGGCAAGTT	9676	QY	3564	SerArgLysArgArgThrGlyPhe	3571
QY	3204	SerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeu	3223	Db	10757	TCCAGGAAAGGAGGACTGGGTTT	10780
Db	9677	TCTGTGTATGTGCAGAGGGTATACCTTTGAGGGAGTTAAACATATCAGTATGTACGTT	9736	RESULT 4			
QY	3224	AspGlyThrTrpGluProPheSerAspGluSerCysSerProValSerCysGlyLys	3243	AX686196			
Db	9737	GATGAACTCTGGAGCCACCATCTCCAGTGAATCTTGAGTCCAGTCTTCTTGCGGAAA	9796	LOCUS	AX686196	11158 bp	DNA
QY	3244	ProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheLysSerThrIle	3263	DEFINITION	Sequence 5 from Patent WO02059315.	linear	PAT 29-MAR-2003
Db	9797	CCTGAAAGTCCAGACATGATTTGTGGTGGCAGTAAATACACCTTTGAAAGCAAT	9856	ACCESSION	AX686196		
QY	3264	IleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGlu	3283	VERSION	AX686196.1	GI:293172028	
Db	9857	ATTATATCAGTGTGAGCTGGCTATGACTAGAGGGGACAGGAGGAGCTGTCTGCCAGGAG	9916	KEYWORDS			
QY	3284	AsnArgGlnTrpSerGlyValAlaIleCysLysGluThrArgCysGluThrProLeu	3303	ORGANISM	Homo sapiens (human)		
Db	9917	AACACACAGTGGAGTGGAGGGGTGGCAATATGCAAGACAGCCAGTGTGAAATCCACTT	9976	REFERENCE	Homo sapiens		
QY	3304	GluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrGlyProAsnValValTyr	3323	AUTHORS	Shinketsu, R.A., Patturajan, M., Vernet, C.A., Casman, S.J., Malyankar, U., Shenoy, S., Spytek, K.A., Gangolli, E., Miller, C., Boldog, F., Li, L., Taupier, R.J., Kekuda, R., Smithson, G., Zerhusen, B.D., Liu, X., Colman, S.D., Tchernev, V., Si, J., Edinger, S., Stone, D., Sciore, P., Millet, I. and Rothenberg, M.		
Db	9977	GAAATTTCTCAATGGGAAGCTGACATTTGAAACAGGACACTGGACCCCACTGGGTATAT	10036	TITLE	Human nucleic acids and polypeptides and methods of use thereof		
QY	3324	SerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGly	3343	JOURNAL	Patent: WO 02059315-A 5 01-AUG-2002;		
Db	10037	TCCTGCACAGAGGCTACAGTCTTGAGGGGCCATCTGAGGCACACTGCACAGAAAATGCA	10096	FEATURES	Curagen Corporation (US)		
QY	3344	ThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPheValIle	3363	source	Location/Qualifiers		
Db	10097	ACCTGGAGCCACCCAGTCCCTCTCTGCAAAACCAATCCATGCCCTGTCTCTTTGTGATT	10156		1. 11158		
QY	3364	ProGluAsnAlaLeuSerGluLysGluPheThrValAspGluAsnValSerIleLys	3383	ORIGIN	/organism="Homo sapiens"		
Db	10157	CCCGAAGATGCTCTGCTGCTGAAAGGAGTGTATGTGTATGATCAGATGTGTCATCAAA	10216	Alignment Scores:	/mol_type="unassigned DNA"		
QY	3384	CysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThr	3403	Pred. No.:	/db_xref="taxon:9606"		
				Score:	0	Length:	11158
				Percent Similarity:	19566.00	Matches:	3503
				Best Local Similarity:	98.57%	Conservative:	16
				Query Match:	98.12%	Mismatches:	47
				DB:	97.96%	Indels:	4
					6	Gaps:	2

US-09-977-053-4 (1-3571) x AX686196 (1-11158)

Qy 4 ArgLeuAlaPheCysCysTyrGlyLeuAlaLeuValSerGlyTyrAlaThrPheGlnGln 23
Db : : : : :
83 AGAATTTGGCGGGCTTGGTGGTCTGGCGCTCGTTGGGCTGGCGACCTTTCAGCAG 142
Qy 24 MetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaProGlyAlaPro 43
Db 143 ATGTCCCGCGTCCGCAATTTTCAGCTTCGCGCTCTTCCCGCGAGCGCGCGCGCGCC 202
Qy 44 GlySerIleProAlaProAlaProGlyValSerGlyAlaAlaGlySerArgValGluArg 63
Db 203 GGGAGTATCCCCCGCGCGCTCTTCCCGCGAGCGCGCGCGCGAGTGGAGCGG 262
Qy 64 LeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGluArgLeuGlu 83
Db 263 CTGGCCAGCGGCTTC-----CGCGTGGCGTCTGGCGAGCTCAGCGAGCGCTGGAG 316
Qy 84 LeuValPheLeuValAspAspSerSerValGlyGluValAlaAsnPheArgSerGluLeu 103
Db 317 CTGTCTTCTGGTGGATGATTTGGTCCAGCGTGGCGGAAGTCAACTTCCGCGAGCGCTC 376
Qy 104 MetPheValArgLeuLeuSerAspPheProValValProThrAlaThrArgValAla 123
Db 377 ATGTTCGTCGCAAGCTGCTGTCCGACTTCCCGTGGTGGCGCGCGCGCTGGCC 436
Qy 124 IleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIleSerThrArg 143
Db 437 ATCGTGACTTCTCGTCCAGAGACTACGTGGTGGCGCGCGTGGATACATCTCCACCGCG 496
Qy 144 ArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIleSerTyrArg 163
Db 497 CGCGCGCGCCACGACAGTGGCGCTGCTCTCCAGAGATCCCTGCCATCTCTCTACCGA 556
Qy 164 GlyGlyGlyThrTyrThrLysGlyValAlaPheGlnGlnAlaGlnIleLeuLeuHisAla 183
Db 557 GGTGGCGGACCTACACCGAGCGCGCTTCCAGCAGCGCGCGCAATTCCTTCTCATGCT 616
Qy 184 ArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsnGlyGlyAsp 203
Db 617 AGAGAAACTCACAAAAGTTGTATTTCTCATCTAGTGGATATTCCTCAATGGCGGAGAC 676
Qy 204 ProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheThrPheGlyIle 223
Db 677 CCTAGACCAATTCGAGCGCTCATCGAGATTCAGAGTGGAGATCTTCACACTTTTGGCATA 736
Qy 224 TrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGluGluHisCys 243
Db 737 TGGCAGGGAACTTCGAGACTGATGACATGCGCTTCACCCCAAGGAGGAGCACTGT 796
Qy 244 TyrLeuLeuHisSerPheGluPheGluAlaLeuAlaArg-----ArgAlaLeuHis 261
Db 797 TACCTGTCTACACAGTTTTGAAGATTTTGGAGCTTTAGTGGCGCTCTGTGCATATGTATT 856
Qy 262 GluAspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyrLeuCys 281
Db 857 GTAGATCTACCTTCTGGGAGTTTTATTTCAGATGATATGGTCCACTGCTCATATCTTTGT 916
Qy 282 AspGluGlyLysAspCysAspArgMetGlySerCysLysCysGlyThrHisThrGly 301
Db 917 GATGAGGCAAGGACTGCTGTGACCGAATGGGAGCTGCANATGTGGGAAACACACAGGC 976
Qy 302 HisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThr 321
Db 977 CATTTTGTAGTGCATCTGTGAAAAGGGGTATAACGGAAAGGCTCTGCAATGATGACTGCACA 1036
Qy 322 AlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCysIle 341
Db 1037 GTTTGGCCCATCGGGACATACAAACCTGAAGGCTCACCGAGGAAATCAGCAGTTTGCATT 1096
Qy 342 ProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysVal 361
Db : : : : :
1097 CCATGTCTGTATGAATAATCACACCTCTCCACTGGAGACCATCCCTGAGACTGTGTCTC 1156
Qy 362 CysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeu 381
Db 1157 TGCAGAGAGGGATACAGGGCATCTGGCCAGACCTGTGAAGTTGTTCACCTGCCCCTGCCCTG 1216
Qy 382 LysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAlaAla 401
Db 1217 AAGCCTCCCGAAAAATGGTTACTTTATCCAAACACTTTCACAAACCACTTTCANTGAGCC 1276
Qy 402 CysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCysLeu 421
Db 1277 TGTGGGCTCCGATCTCACCTGGATTTGATCTGTGGGAAGCAGCATCATCTTATGTCTA 1336
Qy 422 ProAsnGlyLeuThrProSerGlySerGlySerTyrCysArgValArgThrCysProHisLeu 441
Db 1337 CCCAATGGTTTGTGGTCCGGTTCAGAGAGCTACTGCAGAGTAAGAACAATGTCTCATCTC 1396
Qy 442 ArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThrThr 461
Db 1397 CGCCAGCGCAAAACATGGCCCATCAGCTGTCTACAGGGGAAATGTTATATANGAACAA 1456
Qy 462 CysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGln 481
Db 1457 TGTGTGGTGGCTGTGATGAAGGTCACAGGCTAGAGGCGAGTATAGCTTACTTGTCAA 1516
Qy 482 GlyAsnSerGlnThrAspGlyProGluProArgCysValGluArgHisCysSerThrPhe 501
Db 1517 GGAAACAGCCAGTGGGATGGCCAGAACCCCGGTGTGTGGAGCGCACCTGTTCACACCTTT 1576
Qy 502 GlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLysPhe 521
Db 1577 CAGATGCCCAAGATGTCATCATATCCCCCACAACTGTGGCAGCAGCCAGCCAAATTT 1636
Qy 522 GlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGluMet 541
Db 1637 GGGAGATCTGTATGTAAGTTGGCGCAAGGGTTCAATTTATCTGGAGTCAAGAAATG 1696
Qy 542 LeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaAlaValCysLysAsp 561
Db 1697 CTGAGATGTACCACTTCTGGAAAATGGAAATGCGAGTTCAGGCGACCTGTGTGTAAGAC 1756
Qy 562 ValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlnGln 581
Db 1757 GTGAGGCTCTCTCAATCACTGCTTAAAGACATAGAGGCTTAGACTCTGGACAGCNA 1816
Qy 582 AspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyLysVal 601
Db 1817 GATTCGCCAATGTTACTGGCAGATTCACACAGCTAAAGACAACCTCTGGTGAAGAGTG 1876
Qy 602 SerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspValAla 621
Db 1877 TCAGTCCACGTTTCATCCAGCTTTTCCACCCCACTTACCTTTTCCCAATTTGGAGATGTGCT 1936
Qy 622 IleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIleLys 641
Db 1937 ATCGTATACACGGCACTGACCTATCCGCAACCAAGCCAGCTGCAATTTTCCATATCAAG 1996
Qy 642 ValIleAspAlaGluProProValIleAspTyrCysArgSerProProProValGlnVal 661
Db 1997 GTTATTGATGACAGAACCACTCTCATAGACTGGTGCAGATCTCCACCTCCCGTCCAGGTC 2056
Qy 662 SerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGlyVala 681
Db 2057 TCGGAGAGGTACATGCGCAAGCTGGGATGAGGCTCAGTTCTCAGACACTCAGGTGCT 2116
Qy 682 GluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyLysThrIle 701
Db 2117 GAATTTGTTTATTACAGAGTTCATACACAGGAGACCTTTTCCCTCAAGGGGAGACTATA 2176
Qy 702 ValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIleVal 721
Db 2177 GTACAGTATACAGCCACTGACCCCTCAGGCAATAACAGGACATGTGATATCCATATTGTC 2236

QY 722 IleValGlySerProCysGluIleProPheThrProValaenGlyAspPheIleCysThr 741
DB 2237 ATAAAGGTTCTCCCTGTGAAATTCATTCACACCTGTAAATGGGATTTTATGCACT 2296
QY 742 ProAspAsnThrGlyValaenCysThrLeuThrCysLeuGluGlyTyrAspPheThrGlu 761
DB 2297 CCAGATAATACTGGAGTCAACTGTACATTAACCTTGCTTGGAGGGCTATGATTTTCACAGAA 2356
QY 762 GlySerThrAspLysTyrCysAlaTyrGluAspGlyValTyrLysProThrTyrThr 781
DB 2357 GGGTCTACTACCAAGTATTTATGTGCTTATGAGATGGCGTCTGGAACCAACATATACC 2416
QY 782 ThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPheGlu 801
DB 2417 ACTGAATGGCCAGACTGTGCAGTAAGCGTTTGTCAAAACACCGGTTTCAAGTCTTTGAG 2476
QY 802 MetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGluAla 821
DB 2477 ATGTTCTACAAAGCAGCTCGTTGTGATGACACAGATCTGATGAAGAAGTTTCTGAAGCA 2536
QY 822 PheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIleAsp 841
DB 2537 TTTGAGACGACCTGGGAAAAATGGTCCCATCTTTGTAGTGATGCAGAGGACATTCAC 2596
QY 842 CysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyrGlu 861
DB 2597 TGCAGACTGGAGAGAACCTGCACAAAAATATTCCTAGATATAATATGACTATGAA 2656
QY 862 AsnGlyPheAlaIleGlyProGlyGlyTyrGlyValAlaAlaAsnArgLeuAspTyrSerTyr 881
DB 2657 AATGGCTTTGCCAATTTGTCAGGTGGCTGGGGTGCAGCTAATAGCTGGATTAATCTTAC 2716
QY 882 AspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSerSer 901
DB 2717 GATGACTTCTCGACACATGTGCAAGAACAGCCACAGACATCGGCAATGCCAATGCCATCTCA 2776
QY 902 ArgIleLysArgSerAlaProLeuSerAspTyrLysLysLeuIlePheAsnIleThr 921
DB 2777 CGGATTTAAAGAGAGTGGCCCATCTATCTGACTATAAAATTAAGTTAAATTTTAAACATACA 2836
QY 922 AlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGlnArg 941
DB 2837 GCTAGTGTGCCATATCCCGATGAAGAAGATGATACCTTCGATGGGAATATCAGAACGA 2896
QY 942 LeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAspPro 961
DB 2897 CTCCTTCAGACATTTGGAAACTATACAAATAAACTGAAAGAGGACTCTCAACAAAGACCCC 2956
QY 962 MetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGluThr 981
DB 2957 ATGTATTCCTTTTTCAGCTTGCATCAGAAATATCTATAGCCACAGCAATTCATTAGAAACA 3016
QY 982 LysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysValAsn 1001
DB 3017 AAAAAGGCTTCCCTCTTCAGACACGAGCTCAGTGTGAGAGGGCGTATGTGTCAAT 3076
QY 1002 CysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIleGly 1021
DB 3077 TGCCCTTTGGAAACCTATTATATATCTGGAACATTTTCACTGTGAAAGCTGCCGGATCGGA 3136
QY 1022 SerTyrGlnAspGlnGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyrThr 1041
DB 3137 TCTATCAGATGAAAGAGGGCAACTTGAAGTGCAGCTTTTGCCCTCTGGGATGTACAG 3196
QY 1042 GluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThrTyr 1061
DB 3197 GAATATATCCATTCAAGAACATCTCTGATTTGTAAGCTCAGTGTAAACAGGCAAGCCCTAC 3256
QY 1062 SerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLysPhe 1081
DB 3257 TCATACAGTGGACTTGAGACTTGTAATCGTTCACCTGGGCACTTATCAGGCCAAAAATTT 3316

QY 1082 GlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAlaVal 1101
DB 3317 GGTTCOCGGAGCTGCCTCTCTGTGTCAGAAAAACACCTCACTGTGAAAAAGAGGCCCTG 3376
QY 1102 AsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGlyLeuMet 1121
DB 3377 AACATTTCTGCATGTGGAGTTCTGTGTCCAGAGAAAAATCTCGCGTTCCTGGGTAAATG 3436
QY 1122 ProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPheCysLeu 1141
DB 3437 CCTGTCAACCATGTCTCGTGACTATTACCAACTAATGAGGAGAGCCCTTCTGCCTG 3496
QY 1142 AlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGluCysSer 1161
DB 3497 GCCTGTCCCTTTTATGAACTTACCCCATTCGTGTGTTCCAGATCCATCACAGATGTTCA 3556
QY 1162 SerPheSerSerThrPheSerAlaAlaGluGluSerValValProProAlaSerLeuGly 1181
DB 3557 AGTTTATAGTTCAACTTCTCAGCGGACAGAGAAAGTGTGTGCCCCCTCTCTCTTGGCA 3616
QY 1182 HisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPheAsnPro 1201
DB 3617 CATATTAAAGAGGCGCATGAATCAGCAGTCAGCAAGTCATGAATGCTTCTTTAACCCCT 3676
QY 1202 CysHisAsnSerGlyThrCysGlnGlnLeuGluValArgGlyTyrValCysLeuCysProLeu 1221
DB 3677 TGCACACATAGTGGAACTTCCAGCAACTTGGGGTGTGTTATGTTTCTCTGTCCACTT 3736
QY 1222 GlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuProCysLeu 1241
DB 3737 GGAATATACAGCTTAAAGTGTGAAACAGACATCGATGAGTGCAGCCCACTGCTTGCCTC 3796
QY 1242 AsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSerGlyTyr 1261
DB 3797 ARCAATGGAGTTTGTAAAGACCTAGTTTGGGGAATTCATTTGTAGTGCCTTCAAGTTAC 3856
QY 1262 ThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerProCysLeuAsnLys 1281
DB 3857 ACAGTAAGCACTGTGAATTTGAACATCAATGAATGTCACTCAATCAATGTAAGTACAG 3916
QY 1282 GlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPheValGly 1301
DB 3917 GCCACCTGTGTGGATGAATTAATTCATACAGTTGTAAATGTCCAGCCAGGATTTTCAGGC 3976
QY 1302 LeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAlaVal 1321
DB 3977 AAAAGGTGTGAAACAGGTATGTATCAACTCAGTGTGTTATTAATCACTTAATTAATGCAATC 4036
QY 1322 CysGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGlyThrArg 1341
DB 4037 TGTGAGACACAGGTGTGGGGATTTCTGTGCAATGCCACCTGGATTTTGGGTACCGCA 4096
QY 1342 CysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThrCysLys 1361
DB 4097 TGTGGAAGAAACCTCGATGATGTCTCAGTCAGCATTCAGCAAAATGCGAGCTACCTGTAAC 4156
QY 1362 AspGlyAlaAsnSerPheArgCysLeuCysAlaIleGlyPheThrGlySerHisCysGlu 1381
DB 4157 GACGGTGCCCAATAGCTTTCAGGTGCTGTGTGCAAGTGTGAGTGTTCACAGATTCACCTGTGAA 4216
QY 1382 LeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysValAspGlu 1401
DB 4217 TTGACATCAATGAAATGTCACTCTATCCATGTAGAAATCAGGCACCTGTGTGGATGAA 4276
QY 1402 LeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGluThrGlu 1421
DB 4277 TTAATTTATACAGTTGTAAATGTCCAGCCAGGATTTTCAGGCCAAAAAGGTGTGAAACAGAA 4336
QY 1422 GlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrValMetLeu 1441
DB 4337 CAGTCTACAGGCTTTTAACTGGATTTTGAAGTTTCTGGCATCTATGGAATATGTATGCTA 4396
QY 1442 AspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrMetLysSerSerAsp 1461

Db 4397 GATGGCATGCTCCCATCTCTCCATGCTCTAACTGTACCTTCTGGATGAATCCCTCTGAC 4456
QY 1462 AspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsnThrLeu 1481
Db 4457 GACATGAACATATGGAACACCAATCTCTATGAGTTGATTAACGGCAGCACAATACCTTG 4516
QY 1482 LeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAspGlyArgGluLysIleThrAsn 1501
Db 4517 CTCCTGACTGATTAATAACGGGTGGTCTTTATGTGAATGGCAGGAGAAAGATAACAAAC 4576
QY 1502 CysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSerAlaAsn 1521
Db 4577 TGTCCCTCGTGAATGATGCAGATGGCATCATATATGCAATCACTTGGCAAGTACTCGT 4636
QY 1522 GlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeuSerVal 1541
Db 4637 GGAGCCTGGAGGCTCTATATAATAGGGAATATCTGACGGTGTACTGGCCTCTCCAT 4696
QY 1542 GlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGlnAspLysLysGly 1561
Db 4697 GCGAAGCCATACCTGCTGGGTGCATTTAGTTCTTTGGGCAAGAGCAAGCAAAAAGGA 4756
QY 1562 GluGlyPheSerProIleGluSerPheValGlySerIleSerGlnLeuAsnLeuTyrAsp 1581
Db 4757 GAGGGGTTCACCCCGCTGAGTCTTTTGTGGGTCTCATAGCCAGCTCAACCTCTGGGAC 4816
QY 1582 TyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGluLeuSer 1601
Db 4817 TATGTCCTGTCTCCACAGCAGGTGAAGTCACTGGCTACCTCTCTGCCAGAGGAACCTCAGT 4876
QY 1602 LysGlyAsnValIleAlaTyrProAspPheLeuSerGlyIleValGlyLysValLysIle 1621
Db 4877 AAAGGAAACGTGTAGCATGGCTGATTTCTTGTGAGGAATCTGGGAAAGTGAAGATC 4936
QY 1622 AspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValProHisLeu 1641
Db 4937 GATTTCAAGACATATTTTGTCTGATGTCACCGCTTGGGAGGTTCAGTGCCTCATCTG 4996
QY 1642 ArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAspProGly 1661
Db 4997 AGAATGCACTCTGAAGATTTAAACCAAGGTTCGAAAGTCAATCTGTTCTGTGACCAAGC 5056
QY 1662 PheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTyrThrGlnPro 1681
Db 5057 TTCCAGCTGTCGGGAACCTGTGTCAGTACTGTCTGAATCAAGGACAGTGGACACAACCA 5116
QY 1682 LeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGlyPheHis 1701
Db 5117 CTCGCCCACTGTGAACGCATTCGCTGGGGTGCACCTCTTTGGAGAAATGGCTTCCAT 5176
QY 1702 SerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGlyTyrTyr 1721
Db 5177 TCAGCCGATGACTTCTATGCTGGCAGCAGCACTAAGTCACTACAGTGCACCAATGGCTACTAT 5236
QY 1722 LeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTyrAsnGlyValSerPro 1741
Db 5237 CTATTGGGTGACTCAAGGATGTTCTGTACAGATAATGGGAGCTGGAAACGGCTTTCACCA 5296
QY 1742 SerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisIleAspCys 1761
Db 5297 TCTCGTGTAGATGTCGATGAGTGTGAGTGTGATGATGATGATGATGATGATGATGATGAT 5356
QY 1762 LeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAspGlyLys 1781
Db 5357 CTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5416
QY 1782 AsnCysAlaGluProIleLysCysValAlaProGlyAsnProGluAsnGlyHisSerSer 1801
Db 5417 AACTGTGCAGAACCTATAAAATGTAAGGCTCCAGGAAATCCGGAATATGGCCACTCTCA 5476
QY 1802 GlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrGlnLeu 1821

Db 5477 GGTGAGATTTATACAGTAGTCCGGAAGTACACATTTTGTGTGAGGAGGATACCAAGTTG 5536
QY 1822 MetGlyValThrLysIleThrCysLeuGluSerGlyGluTyrAsnHisIleLeuIleProTyr 1841
Db 5537 ATGGGAGTAAACCAATATCAATGTTTGGAGTCTGGAGATGGAAATCATCTAATACCATAT 5596
QY 1842 CysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGluGluLeu 1861
Db 5597 TGTAAAGCTGTTTCATGTGTAAACCGGCTATTTCCAGAAAAATGTTGTCATTGAGGAGTTA 5656
QY 1862 AlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGly 1881
Db 5657 GCATTTTACTTTTGGCAGCAAGTGCATATAGGTGTAATAAAGGATATATCTCTGGCCGGT 5716
QY 1882 AspLysGluSerSerCysLeuAlaAsnSerSerTyrSerHisSerProValCysGlu 1901
Db 5717 GATTAAGATCATCTCTGTTCTTGTACAGTCTTGGAGTCAATTTCCCTCTCTGTGTGAA 5776
QY 1902 ProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSerGlyLeu 1921
Db 5777 CCAGTGAAGTGTCTAGTCCGGAATAATAATAATATGAAAAATATATTTTGTAGTGGGCTT 5836
QY 1922 ThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer 1941
Db 5837 ACCTACCTTTCTACTGTCATCATATTCATGCGATACAGGATACAGCTTTACAGGGCCCTTCC 5896
QY 1942 IleIleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHisLeuVal 1961
Db 5897 ATATTGAATGCACGGCTTCTGGCATCTGGACAGAGGCGCACCTGCTGTCACTCTGTC 5956
QY 1962 PheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPhe 1981
Db 5957 TTCTGTGGAGAACCACTGCCATCAAGATGCTGTCTATTACGGGGAATAAATCTTCACTTTC 6016
QY 1982 ArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIle 2001
Db 6017 AGGAACACCGTCACTTACACTTGCAGAAAGGCTTATCTCTTGTGTGCTTGTACACCAT 6076
QY 2002 GluCysLeuAlaAspGlyLysTyrSerArgSerAspGlnGlnCysLeuAlaValSerCys 2021
Db 6077 GAATGCTGGCCGACCGCAAGTGGAGTAGAAGTGACCAAGCTGCTGGCTGTCTCTCTGT 6136
QY 2022 AspGluProProIleValAlaAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAsp 2041
Db 6137 GATGAGCCACCACTTGTGGACCAACGCTCTCCAGAGACTGCCCATCGGCTTTTGGAGAC 6196
QY 2042 IleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCys 2061
Db 6197 ATTGCATTTCTACTACTGCTCTGATGTTACAGCTAGCAGACAAATTTCCAGCTTCTCTGC 6256
QY 2062 AsnAlaGlnGlyLysTyrValProProGluGlyGlnAspMetProArgCysIleAlaHis 2081
Db 6257 AATGCCAGCGCAAGTGGGTACCCCCAGAAAGGTCAAGACATGCCCGCTGTATAGCTCAT 6316
QY 2082 PheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLysAlaLys 2101
Db 6317 TTCTGTGAAAAACCTCCATCGGTTCCTATAGCATCTTGGAAATCTGTGAGCAAGCAAAA 6376
QY 2102 PheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThrSer 2121
Db 6377 TTTGCACTGGCTCAGTTGTGAGCTTTAATGATGAGAGGCTTTGTACTGAACACCTCA 6436
QY 2122 AlaLysIleGluCysMetArgGlyGlyGlnTyrAsnProSerProMetSerIleGlnCys 2141
Db 6437 GCAAGATTTGAATGTATGAGAGTGGGCGAGTGGAAACCTTCCCCCATGTCCATCCAGTGC 6496
QY 2142 IleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySerAsn 2161
Db 6497 ATCCCTGTGGGTGTGGAGAGCCCAAGCATCATGAATGGCTTATGCAAGTGAATCAAC 6556
QY 2162 TyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyLeu 2181
Db 6557 TACAGTTTGGAGCCATGTGTGGCTTACAGCTGCACAGGGGTTTCTACATCAAGGGGAA 6616

QY 2182 LysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHisPro 2201
DB 6617 AAGAAGAGCACCTGGCGAAGCCACAGGCGAGTGGAGTAGTCTATATCCGACGTGCCACCG 6676
QY 2202 ValSerCysGlyGluProProLysValGluAenGlyPheLeuGluHisThrThrGlyArg 2221
DB 6677 GTATCTTGTGGTGAACCACTAAGGTTGAGATGGCTTTCTGGAGCATACAATGGCAG 6736
QY 2222 IlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerPro 2241
DB 6737 ATCTTTGAGAGTGAAGTGAAGTATCAGTGTAACTCCCGGGCTATTAAGTCAGTCGGAAGCTCT 6796
QY 2242 ValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysValPro 2261
DB 6797 GTATTTGTCTGCGACAGCAATGGCACTGGCAGCAGTGAATCCCTCTGATGTGTCTCT 6856
QY 2262 LeuAspCysGlyLysProProIleGlnAenGlyPheMetLysGlyGluAenPheGlu 2281
DB 6857 CTCGACTGTGGAAACCTCCCGCATCCAGATGGCTTTCATGAAGGAGAAACTTTGAA 6916
QY 2282 ValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSer 2301
DB 6917 GTAGGTCACAGGTTTCAGTTTCTGTAATGAGGTTATGAGCTTGTGGTGAAGTCTCT 6976
QY 2302 TrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetProAla 2321
DB 6977 TGGACATGTGCAAAATCTGGCAAAATGGAATGAAGTCAAAATCAAAAGTGCATGCCCTGCC 7036
QY 2322 LysCysProGluProProLeuLeuGluAenGlnLeuValLeuLysGluLeuThrThrGlu 2341
DB 7037 AAGTGGCCAGAGCCGCCCTCTTTGGAAACCACTAGTATTAAGGAGTTGACACCGAG 7096
QY 2342 ValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSerValLeu 2361
DB 7097 GTAGGAGTGTGACATTTCTCTGTAAGAGGCGCATGCTCTGCAAGCGCCCTCTGCTCTG 7156
QY 2362 LysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysLysLysValLeuCys 2381
DB 7157 AAATGCTTGGCATCCCGACCAATGAATGACTTTCTCCCTGTTTGTAAAGATGTTCTTTGT 7216
QY 2382 ThrProProLeuLeuSerPheGlyValProIleProSerSerAlaLeuHisPheGly 2401
DB 7217 ACCCCACCTCCCTAAATTTCTTTTGGTGTCCCATTTCTTCTCTCTCTCTCTCTCTCTCT 7276
QY 2402 SerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeu 2421
DB 7277 AGTACTGTCAAGTATTTCTGTGTAGTGGGTTTTCTTAAGAGGAAATCTACCCACCTC 7336
QY 2422 CysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGluCysPro 2441
DB 7337 TGCACACCTGATGGCACCTGGAGCTCTCCACTGCCAGAAATGTGTTCCAGTAGAATGTCCC 7396
QY 2442 GlnProGluGluLeuProAsnGlyLysIleAspValGlnGlyLeuAlaTyrLeuSerThr 2461
DB 7397 CAACCTTGAGGAATCCCAATGGAATCATTTGATGTGCAGAGGCTTGGCTATCTCAGCACA 7456
QY 2462 AlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGly 2481
DB 7457 GCTCTCTATACCTGCAAGCCAGGCTTTGAATTTGGTGGAAATACTACCCACCTTTGTGGA 7516
QY 2482 GluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeuLysPro 2501
DB 7517 GAAATATGGTCACTGGCTTGGAGGAAACCAACATGTAAGCCCAATGAGTGCCTGGAACCC 7576
QY 2502 LysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThr 2521
DB 7577 AAGGAGATTTGAATGGCAATTTCTTTACACCGCACTACACTATGACAGACCGTTACC 7636
QY 2522 TyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThr 2541
DB 7637 TACTCTTGCAACCGAGGCTTTGGCTCGAAGTCCCGAGTCCCTTGACCTGTTTAGAGACA 7696

QY 2542 GlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnPro 2561
DB 7697 GGTGATTTGGATGTAGATGCCCATCTTGCATATGCCATCCACTGTGATTTCCTCCCAACCC 7756
QY 2562 IleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSer 2581
DB 7757 ATTGAAATGGTTTTGTAGAGGTGCAGATTACAGCTATGTGCTATTAATCATCTACAGT 7816
QY 2582 CysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTyr 2601
DB 7817 TGTCTCCCTGGGTTTCAGGTGGCTGTGTCATGTCATGTCAGACCTGTGAGAGTTCAGGATG 7876
QY 2602 SerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIleAspPhe 2621
DB 7877 TCAAGTTCCATCCCAACATGTATGCCAATAGACTGTGGCCCTCCCTCTCATATAGATTT 7936
QY 2622 GlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAAspMetMet 2641
DB 7937 GGAGACTGTACTAAACTCAAGATGACCGGATATTTTCAGCAAGAAAGACGACATGATG 7996
QY 2642 GluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLysThrTrp 2661
DB 7997 GAAGTTCCATATGTGACTCTCTCCCTCTTATCATTTTGGGAGCAGTGGCTTAAACCTGG 8056
QY 2662 GluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetVal 2681
DB 8057 GAAATACAAAGAGGCTCTCTGCTACACATTCATCAAACTTCTGTATGTGATGACATGTT 8116
QY 2682 SerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuIleCysGlnGlu 2701
DB 8117 TCATACACCTGTAACTCCAGGATATGAATCTTCTGGGAAACCTGTGCTGATCTGCAGAA 8176
QY 2702 AspGlyThrTrpAsnGlySerAlaProSerCysLysSerIleGluCysAspLeuProThr 2721
DB 8177 GATGGAACTTGGAAATGGCAGTGCACCATCTCTGCAATTTCAATTTGAATGTGACTTGCCTACT 8236
QY 2722 AlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyr 2741
DB 8237 GCTCTCGAAATGGCTTTTTCGCTTTTACAGAGACTAGTGGAGTGGTGTGCTGAGTAT 8296
QY 2742 SerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArg 2761
DB 8297 AGCTGTAAACCTCGACACATTTCTAGCAGGCTCTGACTTAAGCTTTGTCTAGAGAAATAGA 8356
QY 2762 LysTrpSerGlyAlaSerProArgCysGluAlaLysSerCysLysLysProAsnProVal 2781
DB 8357 AAGTGGAGTGGTGCCTCCCAACGCTGTGAAGCCATTTTCATGCACCAAAAGCCAAATCCAGTC 8416
QY 2782 MetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCys 2801
DB 8417 ATGAATGGATCCCATCAAGGAGCACTACACATCTGACACCTGTGACCTGTGTACTATGAGTGT 8476
QY 2802 AspProGlyTyrValLeuAsnGlyThrGluArgThrCysGlnAspAspLysAsnTrp 2821
DB 8477 GACCCCGATATGTGCTGAATGGCAGCTGAGAGGAGAACATGCCAGGATGACAAAACTGG 8536
QY 2822 AspGluAspGluProIleCysLysProValAspCysSerSerProProValSerAlaAsn 2841
DB 8537 GATGAGATGAGCCCATTTTGCATTTCTGTGAGCTGACGTTTCCACCCAGCTCTGAGCCAA 8596
QY 2842 GlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGlu 2861
DB 8597 GGCAGGTGAGAGAGACGAGTACACATTCNAAAAGAGATTGAATACACTTGCATGCAAGAA 8656
QY 2862 GlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGly 2881
DB 8657 GGGTCTCTGCTGAGGAGGACGAGTGGCTTGTCTTGCATTTGTCTGCAATGGAAGTTGGAGTGA 8716
QY 2882 AlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyVal 2901
DB 8717 GCCACTCCCACTGTGTGCTGTGATGTGCCATGTGCCACCCCGCACCACTGGCCCAATGGGGTG 8776
QY 2902 ThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGluGlyTyr 2921

Db	8777	ACGGAAGGCCCTGGACTATGGCTTCAATGAAGGAAGTAACATTCCACTGTCACGAGGCGTAC	8836
Qy	2922	IleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluLeu	2941
Db	8837	ATCTTGACCGGTGCTCCAAACTCACCTGTCACTGTCAGTACAGATGGCACTGGGAGCAGAGATT	8896
Qy	2942	ProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsn	2961
Db	8897	CCTCTGTAAACCAGTCAACTGTGGACCTCTGGAAGATCTTGCCCATGGTTTCCCTTAAT	8956
Qy	2962	GlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeu	2981
Db	8957	GGTTTTTCTTTATTCATGGGGGCAATACACGATATCAGTGTCTTCTGCTGTATAAGCTC	9016
Qy	2982	HisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSer	3001
Db	9017	CATGGAAATTCATCAGAAGGTGCTCTCCAAATGGCTCTGGAGTGGCACTCACCTTCC	9076
Qy	3002	CysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAsp	3021
Db	9077	TGCCTGGCTTGCAGATGTTCCACACCACTGAATTTGAATATGGAACTGTCAATGGGCAGAT	9136
Qy	3022	PheAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeu	3041
Db	9137	TTTGACTGTGGAAAGGCAGCCGGATTCAGTGTCTTCAAAGCTTCTCAAGCTCTTAGGACTT	9196
Qy	3042	SerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHis	3061
Db	9197	TCTGAATATCACTGTGAAGCCGATGGCCAGTGGAGCTCTGGGTCCCCCACTGTGAACAC	9256
Qy	3062	ThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrp	3081
Db	9257	ACTTCTTGTGGTTCCTTCCAAATGATACCAATGCGTTTCATCAGTGAGACCAGCTCTGG	9316
Qy	3082	LysGluAsnValIleThrTrpSerCysArgSerGlyTyrValIleGlnGlySerSerAsp	3101
Db	9317	AAGGAAATATGATAAATCTTACAGCTGCAGGTCTGGATATGTCATACAAAGGCAGTTCAGAT	9376
Qy	3102	LeuIleCysThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSer	3121
Db	9377	CTGATTTGTACAGAAAGGGGTATGAGCCAGCCCTTATTCAGTCTGTGAGCCCTTGTGCC	9436
Qy	3122	CysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGlu	3141
Db	9437	TGTGGTTCCTCCACCGTCTGTGCCAATGCAGTGGCAACTGGAGGGGCACACACTATGA	9496
Qy	3142	SerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPhe	3161
Db	9497	AGTGAAGTGAACCTCAGATGCTCGGAAGGTATACGATGGATACAGATACAGATACATTC	9556
Qy	3162	ThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCys	3181
Db	9557	ACCTGTGAGAAAGATGTGGCTGGTCTCCCTGGAGAGAATCTCTGTCAGTCTCTAAAAAATGT	9616
Qy	3182	ProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArg	3201
Db	9617	CCTCTCCGGAAACATACACATATATCTGTACATGGGACGATTTCACTGTGATAGG	9676
Qy	3202	GlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSerValCys	3221
Db	9677	CAAGTTTCTGTGTCAATGTCAGAAGGGTATACCTTTGAGGAGGTAAACATATCAGTATGT	9736
Qy	3222	GlnLeuAspGlyThrTrpGluProPheSerSerAspGluSerCysSerProValSerCys	3241
Db	9737	CAGCTTGATGGAACCTGGAGGCCACCAATCTCCGATGAATCTTGCAGTCCAGTTTCTGT	9796
Qy	3242	GlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPheGluSer	3261
Db	9797	GGGAAACCTGAAAGTCCAGAACATGGATTTGTGGTGGCAGTAAATACACTTTGAAGC	9856
Qy	3262	ThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCys	3281

9857	DB	ACAAATATTATTCAGTGTGAGCGCTCGCTATGAACCTAGAGGGGAAACAGGGAACGTGCTCTGC	9916
3282	QY	GlnGluAsnArgGlnTrpSerGlyGlyValAlaIleCysValSerGluThrArgCysGluThr	3301
9917	DB	CAGGAGNACAGACAGTGGAGTGGAGGGGTGGCAATATGCAAGAGACACAGGTGTGAACT	9976
3302	QY	ProLeuGluPheLeuAsnGlyIysAlaAspIleGluAsnArgThrGlyProAsnVal	3321
9977	DB	CCACTTGAATTTCTCAATGGGAAGCTGACATTGMAAAACAGGACGACTGGACCCCAACGTG	10036
3322	QY	ValTrpSerCysAsnArgGlyTrpSerLeuGluGlyProSerGluAlaHisCysThrGlu	3341
10037	DB	GPATATTCCTGCAACAGAGGCTACAGTCTTGAAAGGGCCATCTGAGGCACACCTGCACAGAA	10096
3342	QY	AsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPhe	3361
10097	DB	ANTGDACTTGAGGCCACCCAGTCCCTCTCTGCACCAACCAATCATGTGCCCTGTTCCTTTT	10156
3362	QY	ValIleProGluAsnAlaLeuLeuSerGluLysGluPheTrpValAspGlnAsnValSer	3381
10157	DB	GTGATTTCCGAGATGCTCTGCTGCTCTGAAAGAGGATTTTATGTTGATCAGAAATGTGTCC	10216
3382	QY	IleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAsp	3401
10217	DB	ATCAAAATGTAGGGAAGGTTTCTGCTGCAAGGGCCACGGCATCATTAACCTGCAACCCCGAC	10276
3402	QY	GluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHis	3421
10277	DB	GAGACGTGGACACAGCAAGCGCCAAATGTGAAAAATCTCATGTGTGTCCACCAAGCTCAC	10336
3422	QY	ValGluAsnAlaIleAlaArgGlyValHisTrpGlnTrpGlyAspMetIleThrTrpSer	3441
10337	DB	GTAGAAATGCAATGCTCGAGGGCTGATATCATATATGAGACATGATCACTACTCA	10396
3442	QY	CysTrpSerGlyTrpMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThr	3461
10397	DB	TGTTACAGTGGATACATGTTGGAGGGTTCTCTGAGGAGTGTGTTTGTAGAAAAATGGAACA	10456
3462	QY	TrpThrSerProPheCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIle	3481
10457	DB	TGGACATCACCTCCCTATTGTCGAGAGCTGTCTGTGATTTTCCATGTGAGAAATGGGGGATC	10516
3482	QY	CysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGluGlu	3501
10517	DB	TGCCAACGCCCAAGTGTGTTCTGTCCAGAGGGCTGGATGGGGCGCTCTGTGAAGAA	10576
3502	QY	ProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTrpGlnCysAsp	3521
10577	DB	CCAAATCTGATTTCTTCCCTGTCTGAAACGGAGGTGCTGTGTGTGGGCCCTTACACAGTGTAC	10636
3522	QY	CysProGlyTrpThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeu	3541
10637	DB	TGCCCCGCTGGCTGGACGGGGTCTCGCTGTTCATACAGCTGTTTGCCAGTCTCCCTGCTTA	10696
3542	QY	AsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThrGlyHis	3561
10697	DB	ANTGTGGAAATGTGTAGACCAACCCGATGTCTGTCTTCTTCTCTGTGACGGGACAT	10756
3562	QY	AsnCysSerArgLysArgThrGlyPhe	3571
10757	DB	TAAGTGTTCAGGAAAGGAGGACTGGGTTT	10786

RESULT 5
AF206329
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE 1 (bases 1 to 11289)
 AUTHORS Gilges,D., Vinit,M.A., Callebaut,I., Coulombel,L., Cacheux,V.,
 Romeo,P.H. and Vigon,I.
 TITLE Polydom: a secreted protein with pentraxin, complement control
 protein, epidermal growth factor and von Willebrand factor A
 domains
 JOURNAL Biochem. J. 352 Pt 1, 49-59 (2000)
 MEDLINE 20517255
 PUBMED 11062057

REFERENCE 2 (bases 1 to 11289)
 AUTHORS Gilges,D., Vinit,M.A., Callebaut,I., Coulombel,L., Cacheux,V.,
 Romeo,P.H. and Vigon,I.
 TITLE Direct Submission
 JOURNAL Submitted (18-NOV-1999) u474, Inserm, Maternite Port-Royal, 123 bid
 de Port-Royal, Paris 75014, France
 FEATURES
 source
 Location/Qualifiers
 1..11289
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C3H/HeN5ic"
 /db_xref="taxon:10090"
 /map="4b-4c2"
 /cell_line="MS-5 stromal cells"
 221..10924
 /codon_start=1
 /product="polydom protein"
 /protein_id="AAG32160.1"
 /db_xref="GI:11177164"
 /translation="MWSRLAFCCMALALVSGWTFNPQVAPSLNPSFRILFPFASPQALG
 ELAVPASSEAAAGSVKRLGRAPRVRRLRSLGLVFLVDESSVSGQTNFLN
 RLKVRKLLSDPPVSTVAIVTVAIVTSSKNVAVRDVISTRAHQKCALLSRIPA
 IYRCGGTYTGAFQAQILHRSNSTKVLFLIDGYSNGGDRPRTAASLRDGEV
 IFPGIQNIRLNDMASTPKRHCYLLHSFEEFALRALHRLPSGSPQIDMA
 RSYLCRAGKDCDMKACGTHGTQFECEKGYKGLQHECTACPSGTPQDAS
 PGGISTCIPCDVSHTSPPGSDPCDCEBQYRSGQTCFVHCPALKPENGFFIQ
 NTKNHFNAACVRCRPGDLVSGSHLCQPNLWSGTESFCRVCTPHLRPQKHGI
 SCSTAESVNTCLVTCNEGVLEGSTLTLCQGNAMGDPKRCVCHCAVFOKDPKV
 IISPSCKQPARPCQQLSCRGYILSGREVRCATSGKWSAKVQTAVCKDVAPO
 ISCPNDISAKTGQDSNVQWPTAKDNGSKVSHVHFAPTPPFPDVAITY
 TATDSNQASCTFYKIDVPEPVIDWCRSPPIQVVEKEHPASDQSDNSGAE
 LVITSSHTQAGPFGHETVWTATDPSGNRTCDIHIKSGPCEVPFPVNGDIFC
 AODAGVNCSLSCRGYVTEGSTKYKCAPEDGIWRPPYSTENPDCAIKRPAHGFK
 SFEMLYKTRCDMDLFTKFSNAPETTLGNWVPSFCDADDIDCLDLTKLYCIEN
 YNENGPAIGCGWAGNRLDYSDHPLDVQSTFTDVGKARSRIKRTVPLSDPKIQ
 LIFNITASVPEERNDLLEENQRLIKLTETINRLKSTLNKEPMYSFQASSETV
 ADSNLETKAPLFCRPGSVLGRMVCNPLGTSYLSHSTCEKMGYSQDEBQLE
 CKLAPPRTHTYLSRVSSECKACQCKGYSSGLETCESPLGTYQPFGRSCLLC
 PETTVKRGAVDISACGVPCCVBPSPRSLTCCVPCPRDYVPNAGKSFCLACPFY
 TTTLTGTSITDCSSFSSTSAEBESIPLVAPGHSQNKYVSSQVFHECLNCPCHS
 GTCQQLRGYVCLCPGTYGLKCTDIDECSLPCLNGGICEDQVGGTFCRSLGYS
 QICENINECISSCLNGTCTDGLASRYCTCVKGMVGHCTEDVNECQSPCLNNAV
 CKDQVGFSCRCPPGLTRCKNVDECSQCPQNGATCKDQANSFRCCQPAFTGTH
 CELANECQSPNCRQATCDVLNSYCKQPGFSGHRCETEQPSGFLNDFEVSIGY
 XVLADVLPLTHAICAPWKSDDVNTGTPISVALEDDKNTSLTIDYNGWLVNG
 KKLTCPSVNDGIWHIAITWTSGWAVRVYINGLSDGTGLSIGRAIPGGALVL
 QEQQKKEGFNPASFGVSLQNLNDYVLSPOQVLLASCPBELSRGNVLWPDF
 LSGITGVKVSNSMCDSELESGVPHLRPASGNRPKPSKVSFLCDFQPMGNPV
 QYCLNQGGQTPLPCHERICRGLPALENGYSFSAEDFHAGSTVYTYQSGYLLGDSR
 MFTDNGSWNGISPSCLDDECAVSGDSEASCLNTNGSVYSCNPPYTDGDKNCAE
 PVKCAKNPNBNGHSGBIYTVGATVFCDEGHELVGVSTITCLTEMDRLRSPCE
 AISCQVPPVNGVDSAFYTGSKVYVRCGYTLSDGERSACLASGWSHSSPVCE
 LVKCSQPDINNGKYLISGLTVLSIASYSCNGYSLQPSLLECTASGMDRAPSPQ
 LVSCGEPYKXADVATGNSFTGNVATICTREGYTLAGPTDITVQANGKNSNHQCL
 AVSCDEPNVPHASPEAHRLPFTAFYCADGVSLADNSQLICNAQNWVPPAGQAV
 PRCTAHCFEPPHSTLESVSKAPAGSVSVFKMEGFLVNTSAKIECLRGWS
 PSLVSQCIIPVRQGSPPITNGYPSGTNYSGAVVAYSCHKGFYIKGSKSTCEATGQ
 KCAKPNBNGHSGBIYTVGATVFCDEGHELVGVSTITCLTEMDRLRSPCE
 WSKPTPTCHPVSCNPPKVENGFLEHTTGRTPSEARFQCNPNYKAAAGSPVPCOANR
 HWSADAPLSCTPLNCGKPPPLONGELKESPEVSKVOPVNCNEGLVGDNSWTCOKS
 GWSKSPKCPVTKCABPPLLENOLVLKELASVGVMTISCKEHALQPSVLKCLP
 SQWNGSPPIKMWLVCPLPFPVGPASSGALHPGSTVKVLCVDGFFLRGSLTLCQ
 ADSTWSSPLPCEVPCQPEILNGIIRVQGLAYLSTLYTCKRPFELVGNATTLQ
 ENGQWLGKPKMCKPTECEPKEILLNGQFSVSFYGGQITTYTCRGRLEPGRKSLTCL

ETGDMDDPPSCDAIHCSDDQPIENGFEVAGDVRVYGMATIIYSCFPQFQVLGHAMQTC
 EBSGSSSTPCVPIIDCGLPPIHDFGDKTVRGQGHFDORDDMMBVPYLAHPQHLBAT
 AKALENTKESPAASHAFLYGTMSVSCPEGYELLGLPVLIQORDCTWNGTASCSISI
 BCDLPVAPENGFIHTFTTMGSAQYSCRPGLHLESHLRCLQNKQMSGTVPRCRAI
 SCSPNPLWNGSIKDDYSYLVGYTYBCDSGYILNGSKKRTCOENRDMWGHBPMPPIV
 DCSPPVFTNGRVKGEYTFQKBIYVSCREGFLEGARSLICLTNGSWGATSCMPV
 RCPAPPQVNGVADGLDYFKKBAFPLLEGVLAQAPRLTQCSNGTWDAEVPVCKPA
 TCGPPADLPQGPFPNGFSFYHGHIQYQCTGYKLGHNPSRRLPNSGWSGSSPCLCP
 RCSTPIIQGTINATDLCGGKTVQIECPKPKLGLSEITCDANGOMSDVPLCEHAC
 GELPTIPNAIVLEGSLSDESNVTVYSCRPYTMQSSDLICTEKALWSQPYPTCEPLSC
 GPPTVANAVATGEATYESVKLRCLLEGVYMDSDTDTFTQQDQDGHWPVERITCSPKK
 CVPNSMTRIRFHGDDFQVNRQSVSCARGFTHEGWNSTQCPDGTWBPFPDESCTIP
 VYCHGEPGSAHGVNKGHSFGSTIVYQCDPGVKLEGNRERICQENRQMSGVAVCKRE
 NRCTPAPFPNGKAVLENTTSGPSLLFCHRGVTLGSGPEAHCTANGTNHHTPLCKP
 NRCPTVPIPENAVLSEKPYFYQDNVSIKCRGFPKLGNGVITCSDEWTHTNARCE
 KISGPPFPHVENAIARGVITYQDMITTSYISGIMLEGSRLSVCLENGTVSPVCKA
 VCRFPQNGVQCPNACSCPDGMWRLCEBPICILPCLNGGRCAVAPYQDCDPTGWTG
 SCRHTATQCPCLNGGKCIKIPNRCHCLSAWTGDCSRKRRAGL"
 221..271
 467..1015
 /note="Region: VWA domain"
 1055..1342
 /note="Region: similar to ephrin (EPB2) cyst rich domain"
 1349..1903
 /note="Region: 3 CCP domains"
 1904..2392
 /note="Region: 2 HVR domains"
 2393..2587
 /note="Region: CCP domain"
 3212..3697
 /note="Region: similar to thyroglobuline type 2 repeats
 (STT2R)"
 3794..4474
 /note="Region: 6 EGF domains"
 4493..5086
 /note="Region: pentraxin (PTX) domain"
 5102..5452
 /note="Region: 2 CCP domains"
 5450..5569
 /note="Region: EGF domain"
 5576..10618
 /note="Region: 28 CCP domains"
 10619..10900
 /note="Region: 3 EGF domains"
 Alignment Scores:
 Pred. No.: 0 Length: 11289
 Score: 16736.00 Matches: 2920
 Percent Similarity: 89.86% Conservative: 289
 Best Local Similarity: 81.77% Mismatches: 356
 Query Match: 83.79% Indels: 6
 DB: 10 Gaps: 5
 US-09-977-053-4 (1-3571) x AP206329 (1-11289)
 QY 1 MetTpProAqLeuAlaPheCysQysTspGlyLeuAlaLeuValSerGlyTTrAlaThr 20
 Db 221 ATGTGGTCGCGCTGGCCCTTTTGTGCTGCGCTCTGGCTGTGTGCTGAGCTGACCAAC 280
 QY 21 PheGlnGlnMetSerProSerArgAenPheSerPheArgLeuPheProGluThrAlaPro 40
 Db 281 TTCCAGCCGCTGGCCCTTGGCTCAACTTCAGCTTCGCTGCTTCCCGAGGCTCTCCG 340
 QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaGlySer 59
 Db 341 GGGGCTCTGGGCAGACTGGCGGTACTCCCGGTCCAGGAGGAGGAGGAGGAGGAGC 400
 QY 60 ArgValGluArgLeuGlyClnAlaPheArgArgValArgLeuValArgGluLeuSer 79
 Db 401 AANGTGGAGCGCTGGCGCGCGCTTCGAGCCGCTGCGGACATGCGGAGCTCAGC 460
 QY 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPhe 99
 ORIGIN

Db 461 GGCACCTTGGAGCTGCTCTCTCTGTTGGAGAGTGGTCCAGGTGGGCAACCAACTTC 520
Qy 100 ArgSerGluLeuMetPheValArgIysLeuLeuSerAspPheProValValProThrAla 119
Db 521 CTCACAGAGCTCAAGTTCTGTGCGCAAGCTGCTGTCGAGCTTCCCGTGGTCTCCAGCGCC 580
Qy 120 ThrArgValAlaIleValThrPheSerSerIysAsnTyrValValProArgValAspTyr 139
Db 581 ACGCGTGTGGCCATCGTCACTCTCTATCCCAAGAACCAAGTGGTGGCGCGCGTGGATTAC 640
Qy 140 IleSerThrArgAlaArgGlnIleIysCysAlaLeuLeuLeuGlnGluLeuProAla 159
Db 641 ATCTCCACGCGCGCGCACCAACACAAAGTGGCGCTACTCAGCGCGGAGATCCCGGCC 700
Qy 160 IleSerTyrArgGlyGlyGlyThrTyrThrIysGlyAlaPheGlnAlaGlnIle 179
Db 701 ATCACCTACCGCGGTGGTGGCACTATACCAAGGGCGCTTCACGCAAGCGCCGCAAAATC 760
Qy 180 LeuLeuHisAlaArgGluAsnSerThrIysValValPheLeuIleThrAspGlyTyrSer 199
Db 761 CTTCCTCACTCTAGAGAAAACCTCCACCAAGTCATATTTCTCATCCGACGCGCTATTCC 820
Qy 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
Db 821 AATGGCGGAGACCCACAGACTTATTCAGCATCGCTTCGGGATTCGGAGTGGAGATCTTC 880
Qy 220 ThrPheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIys 239
Db 881 ACGTTCCGGATTTGGCAGGGATATTCGGGAATTCGGGAATTCGGGATTCGGAGTGGAGATCTTC 880
Qy 240 GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla 259
Db 941 GAAGAACATTTGCTGCTCCACAGTTTGAAGAAATTTAGAGCTTTAGCTCGCAGGGCG 1000
Qy 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr 279
Db 1001 TTGCATGAAGATCTACTCTCTGGAGTTTATCCAAAGAGGATATGCCCGCTGCTTTAT 1060
Qy 280 LeuCysAspGluGlyIysAspCysCysAspArgMetGlySerCysIysCysGlyThrHis 299
Db 1061 CTCCTGAGGCTGGGAAGACTGCTGTGACAAATGGCCAGCTGCAAAATGTGGGACACAC 1120
Qy 300 ThrGlyHisPheGluCysIleCysGluIysGlyTyrTyrGlyIysGlyLeuGlnTyrGlu 319
Db 1121 ACGGTCAATTTGAATGCATCTGTGAGAGGCTATTTACGGAAAGTCTGCAGCATGAG 1180
Qy 320 CysThrAlaCysProSerGlyThrTyrIysProGluGlySerProGlyIysSerSer 339
Db 1181 TGCACAGCTTGGCCATCAGGACATATAAGCCGGAAGCTTCTCCAGGAGGATCAGCACC 1240
Qy 340 CysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAsp 359
Db 1241 TGCATCCCATGCTCTGAGCTAAGCCACACACTCCCCACCTCGGAAGCACTTCCCTGAAGAC 1300
Qy 360 CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379
Db 1301 TCGGTGTCGGAGAGGATACCCAGAGCTGCGCAGACTGTGAGGTGTCCACTGTCT 1360
Qy 380 AlaLeuIysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
Db 1361 GCCCTGAAGCCTCTGAAAAATGTTTTTATACAAAACACTTGCAAAACCACTTCAAT 1420
Qy 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeu 419
Db 1421 GCCGCTGTGGGGTCCGATGTGCCCGGGCTTTGACCTTGTGGGAAGCAGCATCCATTG 1480
Qy 420 CysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysPro 439
Db 1481 TGTCAACCAATGGTTTGTGTCTGGGACAGAAAGCTTCTGCAGAGTGTGAGACGTGCCCC 1540
Qy 440 HisLeuArgGlnProIysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrIys 459

Db 1541 CACCTCGACAGCCCAACACGCGCACATCAGCTGCTCCACTCGCGGAATGTCTTACAAC 1600
Qy 460 ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspIysLeuThr 479
Db 1601 ACCCTGTGTTGGTTTACCTGCAATGAAGATACAGATTAGAAGGCGACATAGGCTTACC 1660
Qy 480 CysGlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSer 499
Db 1661 TGTCAAGGAAATGCCAGTGGGATGGCCGAGGCCCGGTGTGTAGACGCCATTGTGCC 1720
Qy 500 ThrPheGlnMetProIysAspValIleIleSerProHisAsnCysGlyIysGlnProAla 519
Db 1721 ACCTTCCAGAGCCCAACAGGCGTCATATTCTCCACCAGCTGCGCAAGCAGCTGCC 1780
Qy 520 IysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValIys 539
Db 1781 AGGCTGGGATGACCTGTGCTAAGCTGCCGCCAGGATACATTTTATCCGGGGTCAGA 1840
Qy 540 GluMetLeuArgCysThrThrSerGlyIysTyrAsnValGlyValGlnAlaAlaValCys 559
Db 1841 GAA---GTGAGATGTCCACATCTGGGAGTGGAGTGGCCAAAGTTCAGACAGCTGTGTC 1897
Qy 560 IysAspValGluAlaProGlnIleAsnCysProIysAspIleGluAlaIysThrLeuGlu 579
Db 1898 AAAGATGTGGAGCTCCCAAAATCAGCTGTCCAAATGACATTCAGGCAAGACTGGGGAG 1957
Qy 580 GlnGlnAspSerAlaAsnValThrTyrGlnIleProThrAlaIysAspAsnSerGlyGlu 599
Db 1958 CAGCAGGACTCTGCTAATGTCACTGGCAAGTCCCAACAGCTAAAGACAACTCTGGTGA 2017
Qy 600 IysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp 619
Db 2018 AAGGTGTGAGTCCAGCTCCACCAGCTTTACCCCACTTACTCTTCCCAATTTGGAGAC 2077
Qy 620 ValAlaIleValTyrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
Db 2078 GTGGCCATCACCTACAGGCAACCGACTCATCCGGTAACCAAGCCAGCTGCACTTTCTAC 2137
Qy 640 IleIysValIleAspAlaGluProProValIleIleAspTyrCysArgSerProProVal 659
Db 2138 ATTAAGGTGATTTGATGTGGACCGCTGTGATAGATTGTGGCGGATCTCCACCTCCATC 2197
Qy 660 GlnValSerGluIysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSer 679
Db 2198 CAGGTGTAGAGAGGAGGACCCCTGCAGCTGGGATGAGCTTCTCAGACAACTCC 2257
Qy 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
Db 2258 GGGGCTGAATTTGGTCAATTACCAGCAGTCAACACAAGGCGACATGTTCTCATGGGGA 2317
Qy 700 ThrIleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
Db 2318 ACGGTGGTGTGTGATACAGGCCACTGCCCTCAGGCAACACAGGACCTGTGACATCCAC 2377
Qy 720 IleValIleIysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739
Db 2378 ATTGTCAATAAAGTTCTCCCTGTGAGGTCCCTTCACTTCACTTAAAGCGGAGCTTATC 2437
Qy 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
Db 2438 TGTGCCAGGATGATGTGTGGAGTTAACTGTAGCTGAGCTGCAAGGAGGCTATGATTTTC 2497
Qy 760 ThrGluGlySerThrAspIysTyrTyrCysAlaTyrGluAspGlyValTyrIysProThr 779
Db 2498 ACAGAGGCTCACTGAGAGTACTACTGTGTTTTTGAAGATGTTATCTTGGAGACCA 2557
Qy 780 TyrThrThrGluTyrProAspCysAlaIysIysArgPheAlaAsnHisGlyPheIysSer 799
Db 2558 TACTCTACAGAAATGCCAGACTGTGCTATAAAGCGTTTTTGGCAACCACTGTTTCAAGTCC 2617
Qy 800 PheGluMetPheTyrIysAlaAlaArgCysAspAspThrAspLeuMetIysIysPheSer 819
Db 2618 TTTGAAATGCTATACAAAACCACTCGCTGTGTGATGATGATGATCTGTTTAAAGAGTTTCT 2677

QY 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
DB 2678 CGAGCATTTGAGACTACCTCCGGGAAACATGGTCCCGTCTTTTGAACAGTCTGATGAC 2737
QY 840 IleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
DB 2738 ATTGACTGCAGACTGGAGGAC---CTGACCAAAAATACTGCATCGAGTATAATTACAAC 2794
QY 860 TyrGluAsnGlyPheAlaIleGlyProGlyGlyTyrGlyAlaAlaAsnArgLeuAspTyr 879
DB 2795 TATGAAAATGGCTTTGCAATTGGACAGGAGGTGGGTGCGAGCAACAGGCTGGATTAT 2854
QY 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
DB 2855 TCTTACGATCACTTCCTGATGTTGTACAGGAACAACCCACCGATGTGGGCAAGGCCAGA 2914
QY 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn 919
DB 2915 TCGTCAACGATTAAAGAACTGCTCCATTTGTCTGACCCCAAAATTCAGCTAATTTTAC 2974
QY 920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln 939
DB 2975 ATCACAGCTAGGTGCTCCAGAGGAAAGAAAGATACCTTGAATTTGGAGATCAG 3034
QY 940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959
DB 3035 CAGCGACTCATTAAGACATTGGAACAATCACCAATCGCTGAAAGACCTTTGAATAAA 3094
QY 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
DB 3095 GAGCCCATGTATTTCTTCCAGCTGGCTCGGAACAGTGGTGGCTGACAGCAATTTCCCTC 3154
QY 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
DB 3155 GAAACAGAAAAGGCTTTTCTTCTGACAGCAGGCTCTGTGCTGAGGGGGCGCATGTGT 3214
QY 1000 ValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArg 1019
DB 3215 GTCAACTGCCCTCGGAACCTTCTTACTCTGTGAGCATTCACCTGTGAAAGCTGCCTC 3274
QY 1020 IleGlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMet 1039
DB 3275 ATGGGATCTTACCAAGATGAAGAGGCGCAGCTGGAATGCAAGCTCTGTCCCCCAAGGACT 3334
QY 1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
DB 3335 CACACGGAATACCTCCATTCAAGAAGCGTCTCTGAAATGCAAAGCTCAGTGAAGCAAGGC 3394
QY 1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
DB 3395 ACCTACTCTTCAGTGGGTGGAGACCTCGAATGTGTTCCTGGTACTTATCAACCG 3454
QY 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
DB 3455 GAATTTGGATCCCGAGCTGCTCTCTATGCCAGAAACCAACCAACGGTGAAGAGAGA 3514
QY 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
DB 3515 GCGGTGGACATCTCTGCTTGTGGAGTCCCTCCAGTAGTGAAGAATTTCTCCCGTTCTGGG 3574
QY 1120 LeuMetProCysHisProCysProArgAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
DB 3575 CTAAACCTCTGTACCTTGCCTCGAGACTATTACCAACCCATATGACGAGGAAGTCTTC 3634
QY 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
DB 3635 TGCCTCGCTTGTCCCTTTTATGGAACATAACAATCATCTGGGGCCACGTCATCAAGAC 3694
QY 1160 CysSerSerPheSerThrPheSerAlaAlaGluSerValValProProAlaSer 1179
DB 3695 TGCTCAAGTTTGTAGTCTACTTTTCTCAGCAGCAGAGAAGCATAGTGCCTCTCGTGGCC 3754

QY 1180 LeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPheHisGluCysPhePhe 1199
DB 3755 CTGGACATTTCCAGAACAGATGACAGTACAGTCTTTCACGATGCTTCTTA 3814
QY 1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCys 1219
DB 3815 AACCCCTGGCCACAAACAGTGGACCTCCAAACAGCTGGGGGTGTTATGTTCTGTCTCTGC 3874
QY 1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
DB 3875 CCACCTGGATACACAGCTTAAAGTGTGAACAGATATTGATGAATGACGCTCTCTGCCT 3934
QY 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
DB 3935 TGCTCAATGTGGGAATTTGTAGAGACCAAGTTGGGGGATTCAGTCCGAAATGTTCTATTG 3994
QY 1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu 1279
DB 3995 GGTATTTCAGTCAAAATATGTGAAGAAATATATAATGAGTGTATCTCCAGCCCTTGTCTTA 4054
QY 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
DB 4055 AATAAAGSAACTGCACTCAGCGCTTGGCAAGCTACCGCTGTACTCTGTGAAGAGATAC 4114
QY 1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
DB 4115 ATGGGTGTCATCTGTGAACAGACGTCATGATGATGATGATGATGATGATGATGATGATGAT 4174
QY 1320 AlaValCysGluAspGlnValGlyGlyPheLeuCysLysCysProGlyPheLeuGly 1339
DB 4175 GCAATTTGTAAAGACCACTGAGTGGGGGTCTCATGCAAAATGCCACCCGGAATTTTGGGT 4234
QY 1340 ThrArgCysGlyLysValAspGluCysLeuSerGlnProCysLysAsnGlyValThr 1359
DB 4235 ACTGGTGTGAAGAAATATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4294
QY 1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
DB 4295 TGTAGGATGTTGCCAACAGCTTCAGTGTCAATGTCCAGCAGGCTTCACAGGGACACAC 4354
QY 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
DB 4355 TGTGAATCAACATCAACAGTGTCTAGTCCCAACCCATGTAGGAACCCAGGCCACCTGTGTG 4414
QY 1400 AspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
DB 4415 GATGAATCAACTCACTACAGTGTGAATGTCAACCCAGGATTTTCAGGCCACAGGTGTGAG 4474
QY 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
DB 4475 ACAGAACAGCTTCCGGTTTAAACCTGGATTTTGAAGTTTCTGGCATCTACGGGTACGTC 4534
QY 1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSer 1459
DB 4535 CTGTAGATGAGTGTGTCACACCTCCATGTCATCAACCTCGCATTTCTGGATGAATTC 4594
QY 1460 SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn 1479
DB 4595 TCTGATGTCACTCACTACCGGACGCCCATCTCTATGCACTTGGAGGATGAAGAAGACAC 4654
QY 1480 ThrLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIle 1499
DB 4655 ACCTCCCTCTCACTGATTTACAAACGGCTGGTCTTTATGTGAATGGAAGGAAAGATC 4714
QY 1500 ThrAsnCysProSerValAsnAspGlyArgTyrPheHisIleAlaIleThrThrSer 1519
DB 4715 ACCAACTGCCCTCGCTAAATGATGTCATTTGTCATCATATGCAATCATCATGCAAGT 4774
QY 1520 AlaAsnGlyIleThrLysValTyrIleAspGlyLysLeuSerAspGlyGlyAlaGlyLeu 1539
DB 4775 ACTGGTGGAGCTGGAGGGTCTATATAATGGGAATATATCTGACCGTGTGTACTGGCTC 4834
QY 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAspLys 1559

Db 4835 TCCATTGGCAAGCCATACCTGGTGGCGGTGCATTAGTTCTTGGGCAAGAGCAGACAAA 4894
Qy LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579
Db 4895 AAGGAGAGGGGTTCAACCCGGCTGAGTCTTTGTGGCTCCATAAGCCAGCTCAACCTC 4954
Qy 1580 TrpAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4955 TGGGACTATGTCCTGCTCCACAGCAGGTGAAGTTGCTGGCCAGCTCCCTGGCCAGAGAA 5014
Qy 1600 LeuSerLysGlyAsnValLeuAlaTrpProAspPheLeuSerGlyIleValGlyLysVal 1619
Db 5015 CTGAGTCGGGGAACGTTGATGATGGCCGATTTCTGCTCGGATTCAGAGGATCAGGGAGGGTG 5074
Qy 1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro 1639
Db 5075 AAGGTTGATTCCAGCAGCATGTTCTGCTCTGATGCTCGCTCTTGAAGGATCCGTCGCT 5134
Qy 1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
Db 5135 CACCTGAGACCTGCATCAGGAATCGAAAGCCAGGCTCCAAAGTCAGTCTGTTCTGTGAT 5194
Qy 1660 ProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrpThr 1679
Db 5195 CCGGCTTCCAGATGGTGGGAATCCTGTGCAGTATTGTCTGAACCAAGGGCAGTGGACA 5254
Qy 1680 GlnProLeuProHisCysGluArgIleSerCysGlyValProProProLeuGluAsnGly 1699
Db 5255 CAACCACTCCCCCACTGTGACGCAATTCGCTGTGGGCTGCCCTCCGCTTGGAGNATGCG 5314
Qy 1700 PheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGly 1719
Db 5315 TTCTACTCAGCGGAGCACTTCATCGGCGCAGCAGCGTGACCTATCAGTGCACCACTGGC 5374
Qy 1720 TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrpAsnGlyVal 1739
Db 5375 TACTACTGCTGGGGATTCGGAATGTTCTGCACAGACCAAGGGAGCTGGAAACGGCAAT 5434
Qy 1740 SerProSerCysLeuAspValAspGluCysAlaValAlaGlySerAspCysSerGluHisAla 1759
Db 5435 TCACCATCTGCTCGATGATCGATGATGTCAGTGGCTCGGCTCGGACTGTAGTGAGCAGCC 5494
Qy 1760 SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp 1779
Db 5495 TCCTGCTGAAACACCAAGGATCTTACGTATGCTTCTGTAAACCCACCATACACGGGAGAT 5554
Qy 1780 GlyLysAsnCysAlaGluProIleLysCysValAlaProGlyAsnProGluAsnGlyHis 1799
Db 5555 GGGAAAACTGTGCAGAACCTGTAAATGTGAAGGCTCCAGAAATCCAGAAATGGCCAC 5614
Qy 1800 SerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyr 1819
Db 5615 TCTTCTGGTGAGATTACACCGTGGGTACTGAGTCAGTCACATTTCTCTGTGACGAGGGCAC 5674
Qy 1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAsnHisIleuLe 1839
Db 5675 GAGCTGTGGGAGTGAGCACCATCAGTCTGTTGGAGACTGGCGAGTGGGATCGCCTCAGG 5734
Qy 1840 ProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGlu 1859
Db 5735 CCGTCTGTGAAGCAATTTCTGTGGTGTCCACCTGTTCTGGAATGGTGGTGTGATC 5794
Qy 1860 GluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeu 1879
Db 5795 GGGTCGGCATTCACATACGGCAGTAAGGTGGTGTACAGGTGTGATAAAGGATATATCTTG 5854
Qy 1880 AlaGlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProVal 1899
Db 5855 TCTGGGATGAAGAGTCAGCAGCTGCTGATGTTCTGAGTCTCTGAGTCAATTCCTCTCTGTG 5914
Qy 1900 CysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSer 1919

Db 5915 TCGAGCTAGTGAAGTGTTCAGGCTCGAGCATAAATAACGCGCAATACATCTTAAGT 5974
Qy 1920 GlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGly 1939
Db 5975 GGGCTCACCTACTTCTTATTGTCATCGTACTCTGTGAGAACGATACAGTTTACAGGGC 6034
Qy 1940 ProSerIleIleGluCysThrAlaSerGlyIleTrpAspArgAlaProProAlaCysHis 1959
Db 6035 CCATCCCTCTCTTGAATGACACAGCTTCGGCAGCTGGGACAGAGCGCCACTAGCTCTCAA 6094
Qy 1960 LeuValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPhe 1979
Db 6095 CTTGTCTCTGGGAGAGCTCCATCGTCAAAAGATGCTGTCTCATCTCTGGGAGCACTTC 6154
Qy 1980 ThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAsp 1999
Db 6155 ACTTTTGGGAACACACAGTTGCTTACATGCAAAAGGGGTACACCTCTGCTGGGCTGAC 6214
Qy 2000 ThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaVal 2019
Db 6215 ACCATCGTATGCCAGGCCCAAGCGCAATGGAATTCAGATTAACCCAGTGCCTGGCTGTC 6274
Qy 2020 SerCysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhe 2039
Db 6275 TCTGTGACGAGCCCCCAATGTGGACCAAGCTCTCCAGAGACTGCTCACAGGCTCTTT 6334
Qy 2040 GlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeu 2059
Db 6335 GGAGACACCGGCTTTTACTGTGCGGATGTTACAGCTGGCTGGCTGATTAATTCACAGCTC 6394
Qy 2060 LeuCysAsnAlaGlnGlyLysTrpValProProGluGlyGlnAspMetProArgCysIle 2079
Db 6395 ATCTGCAATGCCCAGGGGAACCTGGGTTCCCCCGCGGGCAGGCTGTGCGGCTGCATA 6454
Qy 2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLys 2099
Db 6455 GCTCACTTCTGTGAAAAACCCCATCTGTTCTCTACAGCATCTTGGAACTCTGTAGCAMA 6514
Qy 2100 AlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsn 2119
Db 6515 GCAAGTTTGCAGCTGGCTCGTGTAGTGTCTCAAGTGCATGGAGGGTTTGTGTCTGAAC 6574
Qy 2120 ThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIle 2139
Db 6575 ACCTCAGCGAAGATTGAATGCTCTGAGAGTGGAGTGGAGCCCTCTCTCCCTCTCGGTC 6634
Qy 2140 GlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGly 2159
Db 6635 CAGTGCATCCCGTGGATGCGGAGAGCTCCCAAGCATGCAATGGCTACCCGAGTGGG 6694
Qy 2160 SerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLys 2179
Db 6695 ACAAACTACAGTTTGGGGCGCTGTGGCTGTACAGCTGCACAGGGATTTCTATATCAAG 6754
Qy 2180 GlyLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCys 2199
Db 6755 GGGGAGAGAAGAGACAGCTGTGAGGCCACAGCAGTGGAGTAAACCCACGCCACTGTC 6814
Qy 2200 HisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThr 2219
Db 6815 CATCTGTGCTCTGTAAAGGAGCCCTTAAGTTGAGACGGCTTCTCTGAGCAGCACCCT 6874
Qy 2220 GlyArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGly 2239
Db 6875 GGCAGGACTTTGAGAGCAAGAGTGTCCAGTGCACCCAGGCTATTAAGGCAGCCGGA 6934
Qy 2240 SerProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCys 2259
Db 6935 AGTCTGTGTTGTTGTTCAGGCAATGCCACTGGCAGCAGCAGCCCTCTCTCTCTGTC 6994
Qy 2260 ValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsn 2279
Db 6995 ACCCTCTCAACTGTGGGAACCCCTCCATTCATCAGATGGCTTTTGAAGAGGAGAAAGC 7054

2280 PheGluValGlySerIysValGlnPhePheCysAsnGluClyTyrGluLeuValGlyAsp 2299
2300 SerSerTrpThrCysGlnIysSerGlyLysTrpAsnLysLysSerAsnProLysCysMet 2319
2320 ProAlaLysCysProGluProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThr 2339
2340 ThrGluValGlyValThrPheSerCysLysGluClyHisValLeuGlnGlyProSer 2359
2360 ValLeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleVal 2379
2380 LeuCysThrProProLeuIleSerPheGlyValProIleProSerSerAlaLeuHis 2399
2400 PheGlySerThrValLysTrpSerCysValGlyGlyPhePheLeuArgGlyAsnSerThr 2419
2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
2440 CysProGlnProGluGluLeuProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeu 2459
2460 SerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrLeu 2479
2480 CysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeu 2499
2500 LysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr 2519
2520 ValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
2540 GluThrGlyAspTrpAspValAlaAspAlaProSerCysAsnAlaIleHisCysAspSerPro 2559
2560 GlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIle 2579
2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
2600 GlyTrpSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIle 2619
2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAsp 2639
2640 MetMetGluValProTyrValThrProHisProTyrHisLeuGlyAlaValAlaLys 2659
2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThr 2679
2680 MetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLysCys 2699
2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
2720 ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
2760 AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
2780 ProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyr 2799
2800 GluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLys 2819
2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
2840 AlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsn 2899
2900 GlyValThrGluClyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu 2919
2920 GlyTyrIleLeuHisGlyValProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
2940 GluIleProLeuCysLysProValAsnCysGlyProGluAspLeuAlaHisGlyPhe 2959
2960 ProAsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyr 2979
2980 LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
2999 GACTTTGGTACTGTAAAGTCAGAGATGGCCAGGACATTTTGTATCAAGAAGATGAC 8134

8135 ATGATGAGAGTCCCATATCTGGCT-----CACCTCAACATTGGAGACACAGCTAAG 8188
2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThr 2679
2680 MetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLysCys 2699
2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
2720 ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
2760 AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
2780 ProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyr 2799
2800 GluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLys 2819
2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
2840 AlaAsnGlyGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrp 2879
2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsn 2899
2900 GlyValThrGluClyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu 2919
2920 GlyTyrIleLeuHisGlyValProLysLeuThrCysGlnSerAspGlyAsnTrpAspAla 2939
2940 GluIleProLeuCysLysProValAsnCysGlyProGluAspLeuAlaHisGlyPhe 2959
2960 ProAsnGlyPheSerPheIleHisGlyHisIleGlnTyrGlnCysPheProGlyTyr 2979
2980 LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSer 2999
2999 GACTTTGGTACTGTAAAGTCAGAGATGGCCAGGACATTTTGTATCAAGAAGATGAC 8134
3000 ProSerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGly 3019

Db 9209 |||||CCATCTCTGCTACCTGTCAGGTGTTCCACCCATCATTTCAACAGGGGAACCATCAACGCA 9268
Qy 3020 ThrAspPheAspCysGlyValAlaAlaArgGlnCysPheLysGlyPheLysLeuLeu 3039
Db 9269 ACTGATTTGGGATGTTGGAAGACGGTCAGATTGAGTGCTTCAAGAGCTTCAGCTGCTT 9328
Qy 3040 GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCys 3059
Db 9329 GGACTTTCTGAATCACCTGTGATGCCAATGCTGAC--GTCCCACTGTGT 9385
Qy 3060 GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer 3079
Db 9386 GAGCACGCTCAGTGGGCGCTCTCCCAACCATACCAACGCAATGTTCTTGGAGGGCAGC 9445
Qy 3080 SerTrpLysGluAsnValIleThrTrpSerCysArgSerGlyTrpValIleGlnGlySer 3099
Db 9446 CTTTGGAGGACATGTTGTAATCTACAGCTGCAGACCTGCTACACCATGCAGTAGT 9505
Qy 3100 SerAspLeuIleCysThrGluLysGlyValTrpSerGlnProTrpValCysGluPro 3119
Db 9506 TCAGATCTGATTTGTACGGAAGAGCATATGAGCCAGCTTACCCCAACGCTGTGAACCC 9565
Qy 3120 LeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThr 3139
Db 9566 CTGTCTGTGGACCCCAACCTAGTCCATGTCAGTGGCAACGAGAGGCTCATACC 9625
Qy 3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTrpThrMetAspThrAsp 3159
Db 9626 TATGAAACAAAGTGAATCAGGTGCTGAGAGGGTATGATGATTCGATACAGT 9685
Qy 3160 ThrPheThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLys 3179
Db 9686 ACAITTCACCTGCCAAGATGCCATTTGGGTCCCTGAAAGAAATCACCTGCCAGTCTTAA 9745
Qy 3180 LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAsnAspPheSerVal 3199
Db 9746 AAATGCCCTGTGCCATCAACATGACACGCACTGTTTTCAGGAGATGACTTCCAGGTG 9805
Qy 3200 AsnArgGlnValSerValSerCysAlaGluGlyTrpThrPheGluGlyValAsnIleSer 3219
Db 9806 AACAGCAAGTTCTGTGTCATGTCAGAGGGTTTACCCACAGAGGAGTGAATGCTCA 9865
Qy 3220 ValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProVal 3239
Db 9866 ACATGCCAGCCGACGATACATGAGGACCACTTTCTGATGAATCTCTGTATCCAGTT 9925
Qy 3240 SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTrpPhe 3259
Db 9926 GTTTGTGGGCATCTCTGAAGCCACGCGATGGCTCCGTGGTTGGCAATGAAGCACAGCTTT 9985
Qy 3260 GluSerThrIleIleTrpGlnCysGluProGlyTrpGluLeuGluGlyAsnArgGluArg 3279
Db 9986 GGAAGCACATGTTTACAGTGTCACCTGCTGCTACAAATAGAGGGGACACAGGAAACA 10045
Qy 3280 ValCysGlnGluAsnArgGlnTrpSerGlyValAlaIleCysLysGluThrArgCys 3299
Db 10046 ATCTGCCAGGACACAGACAGTGGAGTGGAGAGTGGCAGTGTGCAGAGAACAGATGT 10105
Qy 3300 GluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyPro 3319
Db 10106 GAGACTCCAGCTGAGTTTCCCAATGGGAGGGTGTCTTGGAAAAACACCAATCTCGAGCC 10165
Qy 3320 AsnValValTrpSerCysAsnArgGlyTrpSerLeuGluGlyProSerGluAlaHisCys 3339
Db 10166 AGCTTCTGTTTCTGTCACAGAGGCTACACCTTGGAGGGTCCCCCGGAGGACACATGC 10225
Qy 3340 ThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProVal 3359
Db 10226 ACTGCAATGGAACTGGAATCACCTGACTCCCTCTGCAACCAACCAATCCATCCCTGTC 10285
Qy 3360 ProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTrpValAspGlnAsn 3379

Db 10286 CCTTTTGTGATCTCTGAGAACGCCGCTCTTCTGAAAAAGAGTTTTATGTGCACAGAAAT 10345
Qy 3380 ValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsn 3399
Db 10346 GTATCTATCAAGTCAGGGAAGGCTTCTGCTCAAGGCAATGGTGTATCATCAGTGCAGC 10405
Qy 3400 ProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPro 3419
Db 10406 CCTGACGAGACATGACACCAACCAATGCCAGATGTGAAAAAATCTCTGTGGTCTCCA 10465
Qy 3420 AlaHisValGluAsnAlaIleAlaArgGlyValHisTrpGlnTrpGlyAspMetIleThr 3439
Db 10466 AGTCACGTAGAGATGCAATGCTCGAGAGGTGATTTACCACTATGGGACATGATCACC 10525
Qy 3440 TyrSerCysTrpSerGlyTrpMetLeuGluGlyPheLeuArgSerValCysLeuGluAsn 3459
Db 10526 TACTCTGTATTACAGTGGCTACATGCTAGAGGTTCCCTCCGAGATGTTTGGCCATAGAAAAT 10585
Qy 3460 GlyThrTrpThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGly 3479
Db 10586 GGAACATGGACACCATCTCTGTTTGCAGAGCTGTCTGTCCGTCCCATGTCCAGATGA 10645
Qy 3480 GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCys 3499
Db 10646 GTGTCTGTCAACGTCCCAATGCTTGTCTATGCCACGAGCTGGATGGAGGTCTCTGT 10705
Qy 3500 GluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTrpGln 3519
Db 10706 GAAGAGCAATATGATCTACTCCCTGTTTGAATGTTGGGCGCTGTGTGGCCCTTATCAG 10765
Qy 3520 CysAspCysProGlyTrpThrGlySerArgCysHisThrAlaValCysGlnSerPro 3539
Db 10766 TGTGACTCCCCACAGGCTGGAGTGGCTGCCCTGTCTATACAGCTACTTGTTCAGTCCCC 10825
Qy 3540 CysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerTrpThr 3559
Db 10826 TGTTAATGGGGGGAATGATAGACCAACCAATGATGATGATGATGATGATGATGATGAT 10885
Qy 3560 GlyHisAsnCysSerArgLysArgThrGly 3570
Db 10886 GGACATGATTTGTTCCAGGAAGAGAGCGCG 10918

RESULT 6

AR435510 11230 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 3 from patent US 6656707.

DEFINITION AR435510

ACCESSION AR435510

VERSION AR435510.1 GI:40198413

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 11230)

AUTHORS Welch, A. A. and Elliott, G. S.

TITLE C3b/C4b complement receptor-like molecules and uses thereof

JOURNAL Patent: US 6656707-A 3 02-DEC-2003;

FEATURES Location/Qualifiers

1..11230

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 11230
Score: 16707.00 Matches: 2915
Percent Similarity: 89.72% Conservative: 289
Best Local Similarity: 81.63% Mismatches: 361
Query Match: 83.65% Indels: 6
DB: 6 Gaps: 5

US-09-977-053-4 (1-3571) x AR435510 (1-11230)

Qy 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20

Qy 740 CysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPhe 759
Db 2393 TGTCCCGAGGATAGTGTGGAGTTAACTGTAGCTGCAAGAGGCGCTATGATTTTC 2452
Qy 760 ThrGluGlySerThrAspLysTyrCysAlaTyrGluAspGlyValTyrLysProThr 779
Db 2453 ACAGAGGTCCTCCTGAGAGTACTACTGTCTTTGAGATGGTATCTGGAGACCA 2512
Qy 780 TyrThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSer 799
Db 2513 TACTCTACAGATGGCCAGAGCTGTATATAAAGTTTGGCAACCATGGTTTCAAGTCC 2572
Qy 800 PheGluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSer 819
Db 2573 TTTGAATGCTATACAAACCACTCGCTGTGATGACATGGATCTGTTTAAGAAGTTTCT 2632
Qy 820 GluAlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAsp 839
Db 2633 GCAGATTTGAGATTACCTGGGGAACATGTCTCGCTTTTGTAGCATGCTGTATGAC 2692
Qy 840 IleAspCysArgLeuGluGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAsp 859
Db 2693 ATTGACTGCAGACTGGAGGAC--CTGACCAAAATACTGTCATCGAGTATATATCAAC 2749
Qy 860 TyrGluAsnGlyPheAlaIleGlyProGlyTyrGlyAlaAlaAsnArgLeuAspTyr 879
Db 2750 TATGAAATGGCTTTGCAATTTGGACCGAGGCTGGGGTGGAGCAACAGGCTGGATAT 2809
Qy 880 SerTyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLys 899
Db 2810 TCTTACGATCATCTTCCTGGATGTTGTACAGGAACACCCAGGATGGGGCAGCCAGA 2869
Qy 900 SerSerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsn 919
Db 2870 TCGTCAAGGATTAAGAAGACTGTCCCATTTGTCTGACCCCAAAATTCAGCTAAATTTTAA 2929
Qy 920 IleThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln 939
Db 2930 ATCACAGCTAGCGTCCACTCCAGAGGAAGAAACGATACCTTGAATTTGGAGAAATCAG 2989
Qy 940 GlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLys 959
Db 2990 CAGCGACTCATAGACANTTGGAAACATATCCATTCGCTGGAAGACACCTTGAATATA 3049
Qy 960 AspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeu 979
Db 3050 GAGCCCATGTATTTCTTCCAGCTCGCTCGGAAACAGTGGTGGCTGACAGCAATTTCCCTC 3109
Qy 980 GluThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCys 999
Db 3110 GAAACAGAAAGGCTTTTCTTCTTCGAGACCCAGGCTCTGTGTGAGGGGGGCGCATGTGT 3169
Qy 1000 ValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArg 1019
Db 3170 GTCACTGCCCCCTGGGAACTCTTACTCTCTGGAGCATTCACCTGTGAAAGCTGCTC 3229
Qy 1020 IleGlySerTyrGlnAspGluGluGlnLeuGluCysLysLeuCysProSerGlyMet 1039
Db 3230 ATGGGATCTTACCAAGATGAAGAGGCGAGCTGGAAATGCAAGCTCTGTCTGCCCAAGGACT 3289
Qy 1040 TyrThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGly 1059
Db 3290 CAGCGGAATACCTCCCATTAAGAGCGCTCTGTAATGCAAGCTCAGTGTGAAGCAAGGC 3349
Qy 1060 ThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnPro 1079
Db 3350 ACTTACTTCTCCAGTGGCTGGAGACCTCGAATCGTGTCCGCTGGGTACTTATCAACCG 3409
Qy 1080 LysPheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGly 1099
Db 3410 GAATTTGGATCCCGAGCTGCTCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3469

Qy 1100 AlaValAsnIleSerAlaCysGlyValProCysProGluGlyLysPheSerArgSerGly 1119
Db 3470 GCCGTGGACATCTCTGCTTGTGGAGTGCCTCCCGCAGTAGAGAAATTCCTCCGTTCTGGG 3529
Qy 1120 LeuMetProCysHisProCysProAspAspTyrTyrGlnProAsnAlaGlyLysAlaPhe 1139
Db 3530 CTAACACCTGCTTACCTTGGCTTCGAGACTATTAACCAACCAATGCGAGGAGGCTCTTC 3589
Qy 1140 CysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySerArgSerIleThrGlu 1159
Db 3590 TGCCTCGTCTGCTCTTATGGAACTACAACTCATCTGCGGCCCACTGCCATCACAGAC 3649
Qy 1160 CysSerSerPheSerThrPheSerAlaAlaGluSerValProProAlaSer 1179
Db 3650 TGCTCAAGTTTTAGCTCTACTTCTCAGCAGCAGAGAAAGCATAGTGCCTCCCTGGGCC 3709
Qy 1180 LeuGlyHisIleLysLysArgHisGluLysSerSerGlnValPheHisGluCysPhePhe 1199
Db 3710 CCTGGACATTCCTCCAGAACAGTACGAGTACGAGTCTTTCACGAATGCTTCTTA 3769
Qy 1200 AsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGlyTyrValCysLeuCys 1219
Db 3770 AACCCCTGCCAACACAGTGGAACTGCGCAACAGCTTGGGCGTGGTTATGTCTCTCTGC 3829
Qy 1220 ProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGluCysSerProLeuPro 1239
Db 3830 CCACCTGGATACACAGGCTTAAAGTGTGAACAGATATTGATGAATGACGTCTCTGCT 3889
Qy 1240 CysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIleCysGluCysProSer 1259
Db 3890 TGCCTCAATGGTGAATTTGTAGAGACCAAGTGGGGGATTCACGTGCGAATGTTTCAATG 3949
Qy 1260 GlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSerSerSerProCysLeu 1279
Db 3950 GGCTATTTCAGTCAAAATATGTGAAGAAATATAAATGATGATCTCTCCAGCCCTTGTCTTA 4009
Qy 1280 AsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThrCysValLysGlyPhe 1299
Db 4010 AATAAGAAACCTGCTACCTGACCGCTTGGCAGCTACCGCTGCTGCTGTGTGAAAGGATAC 4069
Qy 1300 ValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsnProCysLeuAsnAsn 1319
Db 4070 ATGGTGTGCACCTGTGAACAGACAGCTCAATGAATGCCAGTCAAGCCCTCTTAAACAC 4129
Qy 1320 AlaValCysGluAspGlnValGlyPheLeuCysLysCysProGlyPheLeuGly 1339
Db 4130 GCAGTTTGTAAAGACCAAGTTGGGGGGTCTCTCGTGCAAAATGCCCAACCCCGGATTTTGGGT 4189
Qy 1340 ThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCysLysAsnGlyAlaThr 1359
Db 4190 ACTCGGTGTGAAGAAATATGTGATGATGCTCTCAGTCAAGCATGCCAATATGGAGCCACT 4249
Qy 1360 CysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGlyPheThrGlySerHis 1379
Db 4250 TGTAAAGATGTGTCCACAGCTTCAGGTGTCAATGTCCAGCAGGCTTTCACAGGACACAC 4309
Qy 1380 CysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsnGlnAlaThrCysVal 1399
Db 4310 TGTGAACCTGAACATCAACAGAGTGTCTAGTCCAAACCCGTGTAGGAAACCAAGGCCACTGTGTG 4369
Qy 1400 AspGluLeuAsnSerTyrSerCysValCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db 4370 GATGAACATACTATACAGTTGTAAATGTACGACAGGATTTTCAGGCCACAGGTGTGAG 4429
Qy 1420 ThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGlyIleTyrGlyTyrVal 1439
Db 4430 ACAGACAGCTTCCGGTTTTTAACCTGGAATTTTGAAGTTTCTGCACTCTACGGGTACGTC 4489
Qy 1440 MetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheTyrMetLysSer 1459
Db 4490 CTGCTAGATGGGTGTGCTGCCAACCCCTCAATGCCGTAAACCTGCGCATTTCTGTGATGAATCC 4549
Qy 1460 SerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAspAsnGlySerAspAsn 1479

Db 4550 TCTGATGTCATCACTACGGACGCCCATCTCTATGCACTTGGAGTGCACAAAGACAC 4609
Qy 1480 ThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsnGlyArgGluLysIle 1499
Db 4610 ACCTTCTCTGACTGATTAACACGGCTGGGTCTCTTATGTGAATGAAAGAAAGATC 4669
Qy 1500 ThrAsnCysProSerValAsnAspGlyArgTyrHisIleAlaIleThrTyrThrSer 1519
Db 4670 ACCAATGCCCCCTCCGTAATGATGGCAATTGGCATATATTCATATCATATGACAGAT 4729
Qy 1520 AlaAsnGlyIleTyrLysValTyrIleAspGlyLysLeuSerAspGlyAlaGlyLeu 1539
Db 4730 ATTGGTGGAGCTGGAGGCTATATAGATGGGAATATCTGACGGTGTGTGGCTC 4789
Qy 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGluAsnLys 1559
Db 4790 TCCATTGGCAAGACCATACCTGCTGGCGGTGCTATTAGTTCTTGGGCAAGACGACAA 4849
Qy 1560 LysGlyGluGlyPheSerProAlaGluSerPheValGlySerIleSerGlnLeuAsnLeu 1579
Db 4850 AAGGAGAGAGGGGTTCACCCGCTGAGTCTTTTGTGGGCTCCATAAGCCAGCTCAACCTC 4909
Qy 1580 TrrAspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4910 TGGGACTATGTCCTGCTCCACAGCAGGTGAAGTTGCTGGCCAGCTCTGCCAGAGGA 4969
Qy 1600 LeuSerLysGlyAsnValLeuAlaTrrProAspPheLeuSerGlyIleValGlyLysVal 1619
Db 4970 CTGAGTGGGGAAACCGTTAGCATGCCCGATTTCTCTGTGGGAATCAACGGGAAGGTG 5029
Qy 1620 LysIleAspSerLysSerIlePheCysSerAspCysProArgLeuGlyGlySerValPro 1639
Db 5030 AAGTTGATTCCAGCAGCATGTTCTGCTCTGATGTGCTGCTTTAGAGGATCCGTGCCT 5089
Qy 1640 HisLeuArgThrAlaSerGluAspLeuLysProGlySerLysValAsnLeuPheCysAsp 1659
Db 5090 CACCTGAGACCTGCATCAGGAATCCGAAGCCAGGCTCCAAAGTCAGTCTGTCTGTGAT 5149
Qy 1660 ProGlyPheGlnLeuValGlyAsnProValGlnTyrCysLeuAsnGlnGlyGlnTrrThr 1679
Db 5150 CCGGGCTTCAGATGTTGGGAATCTCTGTGCAATTTGCTGAACCAAGGGCAGTGGACA 5209
Qy 1680 GlnProLeuProHisCysGluArgIleSerCysGlyValProProLeuGluAsnGly 1699
Db 5210 CAACCACTCCCCACTGTGAACGCATTCGCTGGGGCTGCTCCCTCCCTTGGGAATGGC 5269
Qy 1700 PheHisSerAlaAspAspPheTyrAlaGlySerThrValThrTyrGlnCysAsnAsnGly 1719
Db 5270 TTCTACTCAGCCGAGGACTTCATATGGGGCAGCAGCGTGCACCTATCAGTGCACCATGGC 5329
Qy 1720 TyrTyrLeuLeuGlyAspSerArgMetPheCysThrAspAsnGlySerTrrAsnGlyVal 1739
Db 5330 TACTACTCTGCTGGGTGATTTCCGAAATGTTCTGCRACAGAACCGGGAGCTGGAAACGGCAT 5389
Qy 1740 SerProSerCysLeuAspValAspGluCysAlaValGlySerAspCysSerGluHisAla 1759
Db 5390 TCACCATCTCTGATGTCATGATGATGTCAGTGCAGTCCGCTCGGACTGTAGTGGACAGCC 5449
Qy 1760 SerCysLeuAsnValAspGlySerTyrIleCysSerCysValProProTyrThrGlyAsp 1779
Db 5450 TCTGCTGCAACACCAACGATCTTACGTATGCTCTCTGTAAACCCACCATACACGGGAGAT 5509
Qy 1780 GlyLysAsnCysAlaGluProIleLysCysLysAlaProGlyAsnProGluAsnGlyHis 1799
Db 5510 GGGAAAAATGTGCGAACTCTGTAATATGATAGGCTCCAGAAAAATCCAGAAAAATGGCCGC 5569
Qy 1800 SerSerGlyGluIleTrrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyr 1819
Db 5570 TCTTCTGGGAGATTTACACCTGGGTGACTGACGTGCATTTCTTGTGACGAGGGCAC 5629
Qy 1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGlnTrrAsnHisLeuIle 1839
Db 5669

Db 5630 GAGCTGTTGGAGTGCAGCACCATCACGTGTTTGGAGAGTGGGAGTGGGATCGCTCAGG 5689
Qy 1840 ProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAsnGlyCysIleGlu 1859
Db 5690 CCGTCTCTGTGAAGCCATTTCTGTGGTGTCCACCTGTTCTTCTGAAATGGGTGTGCAC 5749
Qy 1860 GluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeu 1879
Db 5750 GGTTCGGCATTCATATGCGCAGTGAAGTGTGTACAGGTGTGTATAAAGGATATCTTTG 5809
Qy 1880 AlaGlyAspLysGluSerSerCysLeuAlaAsnSerSerTrrPheHisSerProProVal 1899
Db 5810 TCTGGGATGAAGAGTCAGCATGCTTGTGTAGTGTTCCTGGAGTCAATCTCTCTCTGTG 5869
Qy 1900 CysGluProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrIleLeuSer 1919
Db 5870 TGGGGCTAGTGAAGTGTTCGCCAGCTGCAGACATAAATAACGGCAATATCATCTTAAGT 5929
Qy 1920 GlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAspThrGlyTyrSerLeuGlnGly 1939
Db 5930 GGGCTCACCTACCTTCTATTGCATGCTACTCTGTGAGAACGATACAGTTTACAGGGC 5989
Qy 1940 ProSerIleIleGluCysThrAlaSerGlyIleTrrAspArgAlaProProAlaCysHis 1959
Db 5990 CCATCCCTCTGATGACAGCTTCGGCAGCTGGGACAGAGCGCCACCTAGCTGTCAA 6049
Qy 1960 LeuValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPhe 1979
Db 6050 CTGTGCTCTGCGGAGAGCTCCAAATGCTCAAGATGCTGTCACTCATCTGGGAGCACTTC 6109
Qy 1980 ThrPheArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAsp 1999
Db 6110 ACTTTTGGGACACAGTTGCTTACACATGCAAGAGGGCTACACCTTGTCTGGGCTGAC 6169
Qy 2000 ThrIleGluCysLeuAlaAspGlyLysTrrPheSerArgSerAspGlnGlnCysLeuAlaVal 2019
Db 6170 ACCATCATATGCCAGGCCAACCGCAATGCAATTTCAAGTAACACCACTAGTGTCTGGTGC 6229
Qy 2020 SerCysAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPhe 2039
Db 6230 TCTGTGAGAGAGCCCCCAATGAGGACCGCTCTCCAGAGACTGCTCACAGGCTCTTT 6289
Qy 2040 GlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeu 2059
Db 6290 GGAGACACCGGTTTACTTCTGTGGAGTGGCTACAGCTGTGCTGATTAATTTCCAGCTC 6349
Qy 2060 LeuCysAsnAlaGlnGlyLysTrrValProProGluGlyGlnAspMetProArgCysIle 2079
Db 6350 ATCTGCAATGCCAGGGGAACCTGGGTTTCCCGCGCGCCAGGCTGTGCGCGCTGCATA 6409
Qy 2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGluSerValSerLys 2099
Db 6410 GCTCACTCTGTGAAACACCCCATCTGTTCCTACAGCATCTTGGAAATCTGTGAGCAA 6469
Qy 2100 AlaLysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsn 2119
Db 6470 GCAAGTTTGCAGCTGGCTCGGTAGTGAAGTTCAGTGCATGAGGGGTTTGTGCTGAAC 6529
Qy 2120 ThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrrAsnProSerProMetSerIle 2139
Db 6530 ACCTCAGCGAAGATTTGAATGCTCTGAGAGGTGGAGAGTGGAGGCTTCTCCCTCTCGGTC 6589
Qy 2140 GlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGly 2159
Db 6590 CAGTGCATCTCGGTGGGATGCGGAGAGCTTCCAAAGCATTCGAATGCTACCCGAGTGGG 6649
Qy 2160 SerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLys 2179
Db 6650 ACAAACTACAGTTTGGGCGCGTGTGGCTACAGCTGCCCAAGGGATCTATATCAAG 6709
Qy 2180 GlyGluLysIleSerThrCysGluAlaThrGlyGlnTrrPheSerProIleProThrCys 2199
Db 6710 GGGGAGAGAGAGCAGTGTGAGGCGCACAGGACAGTGGAGTAAACCCACGCCCCACCTGC 6769

QY 2200 HisProValSerGlyGluProProLysValGluAenGlyPheLeuGluHisThrThr 2219
DB CATCTGTGTCGTAAACAGCCACTAAGGTGTGAAGCGCTTCTGGAGCACCACT 6829
QY 2220 GlyArgIlePheGluSerGluValArgTyrGlnCysAenProGlyTyrLysValGly 2239
DB GGCAGGACCTTTGAGAGCGAGCAAGGTTCAGTGCACACCCAGGCTATTAAGCGACCGGA 6889
QY 2240 SerProValPheValCysGlnAlaAenArgHisThrHisSerGluSerProLeuMetCys 2259
DB AGTCTGTGTGTGTGCGAAGCAATCGCACTGGCAGCAGCGCCCTCTGTGCTGCG 6949
QY 2260 ValProLeuAenCysGlyLysProProLleGlnAenGlyPheMetLysGlyGluAen 2279
DB ACCCTCTCAACTGTGGGAACCCCTCCATTCAGATGGCTTTTGAAGGAGAAAGC 7009
QY 2280 PheGluValGlySerLysValGlnPhePheCysAenGluGlyTyrGluLeuValGlyAsp 2299
DB TTTGAAGTAGGTCGAAGTTTCAGTTTGTCTGTATGAGGATATAGCTCGTTGGTGAT 7069
QY 2300 SerSerTrpThrCysGlnLysSerGlyLysTrpAenLysLysSerAenProLysCysMet 2319
DB AATCTTGCACTTGCCAGAAATCTGCCAAATCGAGTAAGCAAGCCAGCGAGTGTGTC 7129
QY 2320 ProAlaLysCysProGluProLeuLeuGluAenGlnLeuValLysGlyLeuThr 2339
DB CCCACCAAGTGTGAGAGCCCTCTCTTAGAAAAACCCAGCTCGTATTAAGGAAATAGCT 7189
QY 2340 ThrGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSer 2359
DB TCCGAGGTAGGNGATGACCAATTCCTGTAAAGAGGGCGATGCCCTTGACGGCCCTCT 7249
QY 2360 ValLeuLysCysLeuProSerGlnTrpAenAspSerPheProValCysLysIleVal 2379
DB GTCTCTGAAGTGTGTCATCCGCGCAATGGAATGTTCTTCTTCTATTTGTAAGATGTC 7309
QY 2380 LeuCysThrProProLeuLysSerPheGlyValProLleProSerSerAlaLeuHis 2399
DB CTTTGTCCCTCGCTTCTTGATTCCTTGGCGTCCCTGCGCTTCTGCGTCTTCTAT 7369
QY 2400 PheGlySerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAenSerThr 2419
DB TTTGGCAGTACTGTCAAGTATCTGTGTGTGAGCGGTTTCTTAAGAGCGAGTCCAAC 7429
QY 2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
DB ATCTCTGCGCAGGCTGATAGCACCTGGAGTCTCCATTTGCCGAATGCGTTCCGGTAGAA 7489
QY 2440 CysProGlnProGluGluLeuProAenGlyIleLeuAspValGlnGlyLeuAlaTyrLeu 2459
DB TGTCCCAACCTTGAGAGATCTCAACGATATCATCCAGTAAAGGGCTTGCCTATCTC 7549
QY 2460 SerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAenThrThrThrLeu 2479
DB AGCACCACTCTACACCTGCAAGCCAGGCTTGTAGTTAGTGGCAATGCTACCACCTC 7609
QY 2480 CysGlyGluAenGlyHisTrpLeuGlyLysProThrCysLysAlaIleGluCysLeu 2499
DB TGTGGGAAAAATGGCCAGTGGCTCGGAGGAAACCAATGTGCAAAACCCATTGAATGCCCA 7669
QY 2500 LysProLysGluIleLeuAenGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr 2519
DB GAGCCCAAGGAGATTTAAATGGCAATCTCTCCGTGAGCTTTCAGTATGGCAACCC 7729
QY 2520 ValThrTyrSerCysAenArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
DB ATCACAATCTTTGTGACCGGGCTTCCGGCTCGAAGGTCCCAAAATCCCTGACCTGTTTA 7789
QY 2540 GluThrGlyAspTrpAspValAspAlaProSerCysAenAlaIleHisCysAspSerPro 2559
DB GAGACAGGTGATGGGATATGGATCCCCCTCTTGTGATGCCATCCACTGCAGTACCCA 7849

QY 2560 GlnProIleGluAenGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleIle 2579
DB CAGCCCATTTGAATGGTTTGTAGAGGTGGGATTACAGATACCGTGGCATGATCATC 7909
QY 2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer 2599
DB TATAGCTGCTTCCCTGGGTTTCAAGGTGCTTGGTCATGCCATGCCAGACCTGTGAAGAGTCG 7969
QY 2600 GlyTrpSerSerSerIleProThrCysMetProLleAspCysGlyLeuProProHisIle 2619
DB GGATGGTCAAGCTCCAGCCCAACCTGTGTACCATAGATGCGGCTCTCCCTCTCTCACA 8029
QY 2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGlnGlnGluAspAsp 2639
DB GACTTTGTGTGACTGTACTAAGTCAGAGATGCCAGGACATTTTCATCAAGAGATGAC 8089
QY 2640 MetMetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLys 2659
DB ATGATGGAAGTCCCATATCTGGCT-----CACCTTCAACATTTGGAAGCAACAGCTAAG 8143
QY 2660 ThrTrpGluAenThrLysGluSerProAlaThrHisSerSerAenPheLeuTyrGlyThr 2679
DB GCCTTTGGAAATATCAAAAGGAGTGGCTGCTGCATGCATGCCACTCTCTATGGCAG 8203
QY 2680 MetValSerTyrThrCysAenProGlyTyrGluLeuLeuGlyAenProValLeuLys 2699
DB ATGGTTTCTACGCTCGAGCTGCTGTTATGAACTGCTGGAAATCCCTGTGCTGATCTGC 8263
QY 2700 GlnGluAspGlyThrTrpAenGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
DB CAGGAAGATGTTACGTGGAATGTTACCGCACCTCTTGCATTTCCATTGAATGATTTG 8323
QY 2720 ProThrAlaProGluAenGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
DB CCTGTGTCTCCGAAATGCGCTTTTACATTTACACAGACGACTATGGCAGTGTGCA 8383
QY 2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
DB CAATATAGTGCACACGGGGCACATCTTAGAAGGCTCCACCTTAAGACTCTCTGTGCAG 8443
QY 2760 AsnArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
DB AATAAGCAGTGGAGTGGCACCTGTTCCACGCTGTGAAGCCATCTCATGTCAGTAAAGCCAAAC 8503
QY 2780 ProValMetAenGlySerIleLysGlySerAenTyrThrTyrLeuSerThrLeuTyrTyr 2799
DB CCACCTCTGGAATGATTCATCAAGAGATGACTCTCTCTCTGCGGTGTGTATACTAC 8563
QY 2800 GluCysAspProGlyTyrValLeuAenGlyThrGluArgArgThrCysGlnAspAspLys 2819
DB GAGTGTGACTCTGGCTATATTTCTCAATGGCTCTAAGAGAGGACATGCCAAGAAATAGA 8623
QY 2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839
DB GATTGGGATGGGCATGAGCCCATGTATTCCTGTAGACTGTGGCTCACCCTCCAGTCCCC 8683
QY 2840 AlaAenGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCys 2859
DB ACCAATGGCCGAGTGAAGGAGAGAGATACATTTCAAGAGAGATTTACATCTCTTGC 8743
QY 2860 AsnGluGlyPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAenGlySerTrp 2879
DB COTGAAGGTTTCACTTGAAGGAGCCAGGAGTGTATCTGTCTTACCAATGCAAGTTGG 8803
QY 2880 SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAen 2899
DB AGTGGTGCCACTCCCGAGCTGCATGCTGTGTAGATGTCTGCTGCCCCACACAGAGTGCCAAAT 8863
QY 2900 GlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu 2919
DB GGGGTGGCGAGTGGCTTAGACTATGGTTTCAAGAGAGAGTACGCTTCCACTCTCTAGAG 8923
QY 2920 GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAenTrpAspAla 2939

8924 GGTATGTGTCAGGGGGCTCCAGACTCAGCTGTCAGTCCATGGAGCTTGGGATGCA 8983
QY
2940 GluIleProLeuGlySerProValAsnGlyGlyProProGluAspLeuAlaHisGlyPhe 2959
DB
8984 GAAGTCCCTGCTGTAACACCGAGCTACTGCTGGTCTCTCTCCGCGACCTTCCAGGAGCTTC 9043
QY
2960 ProAsnGlyPheSerPheHisGlyGlyHisIleGlnTyrGlnCysPheProGlyTyr 2979
DB
9044 CCTAATGGCTTTCTTTTATCATGCGGGCCACATACAGTATCAGTGTATTCTGCTGTAT 9103
QY
2980 LysLeuHisGlyAsnSerSerArgArgCysLeuSerAsnGlySerTyrSerGlySerSer 2999
DB
9104 AGCTTCATCGAAGAACCCATCAAGAGATGCTCTCCCATGGTCTCTGGAGCGGAGCTCG 9163
QY
3000 ProSerCysLeuProCysArgCysSerThrProValIleGluTyrGlyThrValAsnGly 3019
DB
9164 CCATCTGCTACTTGCAGGTGTTCCACACCCATCAATCAACAGGGAACCATCAACGCA 9223
QY
3020 ThrAspPheAspCysGlyValAlaAlaArgIleGlnCysPheHisGlyPheHisLeu 3039
DB
9224 ACTGATTGGGAGTGGGAAGACGGTCCAGATTGAGTGTCTCAAGGCTTCAAGCTGCTT 9283
QY
3040 GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTyrSerSerGlyPheProHisCys 3059
DB
9284 GGCATTTCTGAATACCTCTGTGATGCAATGGCCATGGTCTGAC---GTCCCACTGTGT 9340
QY
3060 GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer 3079
DB
9341 GAGCAGCTCAGTGGGGCTCTCCCAACCATACCCCAACGCAATGTGCTTGGGGCAGC 9400
QY
3080 SerTrpLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySer 3099
DB
9401 CTTTTCGAGGACAAATGTGTAACTTACAGCTGAGACCTGGCTACACCATGCAAGGTAGT 9460
QY
3100 SerAspLeuIleCysThrGluLysGlyValTyrSerGlnProTyrProValCysGluPro 3119
DB
9461 TCAGATCTGATTGTACGGAAAGGATATGAGCAGCTTACCCCAAGTGTGAACCC 9520
QY
3120 LeuSerCysGlySerProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThr 3139
DB
9521 CTGTCTGTGGACCCCAACCACTAGCCATGAGTGGCAACAGGAGGCTCATACC 9580
QY
3140 TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp 3159
DB
9581 TATGAAGCAAGTGAATCTCAGGTGCTGGAGGGGTATGTGATGATTCGGATACAGAT 9640
QY
3160 ThrPheThrCysGlnLysAspGlyArgTyrPheProGluArgIleSerCysSerProLys 3179
DB
9641 ACATTCACCTGCGCGCAAGATGGCCATTTGGTCCCTGGAAGATCACCCTGCAGTCCCTAA 9700
QY
3180 LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerVal 3199
DB
9701 AAATGCCCTGTGCTCATCCATCAATGAGGACCCATTTCTGATGATCTGTATCCAGT 9760
QY
3200 AsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAlaIleSer 3219
DB
9761 AACAGACAAGATTCTGTGTCATGTCAGAGGGGTATACCCCAAGAGGATGAATGCTGCA 9820
QY
3220 ValCysGlnLeuAspGlyThrTrpGluProPheSerAspGluSerCysSerProVal 3239
DB
9821 ACATGCCACCCCGAGCTATCGGAGCCACCATTTCTGATGATCTGTATCCAGT 9880
QY
3240 SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPhe 3259
DB
9881 GTTTGTGGGCATCTCGAAGCCCGAGCGATGGCTCCGTGGTGGCAATAGCAGCTTT 9940
QY
3260 GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg 3279
DB
9941 GGAAGCACCATTTGTTTACAGTGTGACCTTGGCTTACAAATATAGAGGGGAACAGGAAACGA 10000
QY
3280 ValCysGlnGluAsnArgGlnTyrSerGlyValAlaIleCysLysGluThrArgCys 3299
DB

10001 ATCTGCCAGGAGAACAGACAGCTGAGTGGAGAGCTGGCAGTGTGCAGAGAGAACAGATGT 10060
QY
3300 GluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrThrGlyPro 3319
DB
10061 GAGACTCCAGCTGAGTTTCCCAATGGGAGGCTGTCTTGGAAACACACACATCTGGAGCC 10120
QY
3320 AsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCys 3339
DB
10121 AGCTTCTGTTTCTCTGTACAGAGGCTACACCTGAGAGGCTCCCGAGGAGCACATGTC 10180
QY
3340 ThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsnProCysProVal 3359
DB
10181 ACTGCAATGGAACTGGAATCACTGACTCCCTCTGCAAAACCAATTCATGCTGCTGTC 10240
QY
3360 ProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsn 3379
DB
10241 CCTTTTGTGATTCTCTGAGAACCGCTCTTCTGANAAGAGTTTATGTGCAGCAGAAAT 10300
QY
3380 ValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsn 3399
DB
10301 GTATCTATCAATGTCAGGAGGCTTCTGCTCAAGGCAATGGTGTATCATCAGTGCAGC 10360
QY
3400 ProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPro 3419
DB
10361 CCTGACGAGACATGAGCAGCACCAATGTCAGATGTGAAAAATCTCTGTGCTCTCCA 10420
QY
3420 AlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThr 3439
DB
10421 AGTCACGTCAGAGTAATGCAATGCTCGAGGAGTGTATTACCAAGTATGGGAGCATGATCACC 10480
QY
3440 TyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsn 3459
DB
10481 TACTCTGTTTACAGTGGCTACATGCTAGAGGTTCCCTCCGAGTGTGCTGAGGAAAT 10540
QY
3460 GlyThrTrpThrSerProProIleCysArgAlaValCysArgPheProCysGlnAsnGly 3479
DB
10541 GGAACATGGACACCATCTCTGTTTGCAGAGCTGCTGTGCTGCTCCCATGTGAGATGGA 10600
QY
3480 GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCys 3499
DB
10601 GGTGTCTGTCAACGTCCTCAAAATGCTTGTCTATGCCAGACGCTCGATGGAGCTCTGT 10660
QY
3500 GluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGln 3519
DB
10661 GAGAGCCAAATGATCATCTCCCTGTTGATGTTGGTGGGCTGTGTTGGCCCTTATCAG 10720
QY
3520 CysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaValCysGlnSerPro 3539
DB
10721 TGTGACTGCCCCACAGGCTGGACTGGGTCCGCTGTATACAGCTACTTGTGAGTCCGCC 10780
QY
3540 CysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTrpThr 3559
DB
10781 TGTCTAAATGGCGGAAATGCAATGAACCAACCAACGATGGCAATGTCTCTCAGCTGGACA 10840
QY
3560 GlyHisAsnCysSerArgLysArgArgThrGly 3570
DB
10841 GGACATGATTGTTCCAGAAAGAGAGAGCCGGG 10873

RESULT 7

AX375163 11230 bp DNA linear PAT 01-MAR-2002
LOCUS
DEFINITION Sequence 3 from Patent WO0210388.
AX375163
ACCESSION AX375163.1 GI:19169917
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE
1
AUTHORS Welch, A.A. and Elliot, G.S.
TITLE C3b/c4b complement receptor-like molecules and uses thereof
JOURNAL Patent: WO 0210388-A 3 07-FEB-2002;

FEATURES		Amgen Inc. (US)		Location/Qualifiers	
source		1. 11230		/organism="Mus musculus"	
				/mol_type="unassigned DNA"	
				/db_xref="taxon:10090"	
ORIGIN					
Alignment Scores:					
Pred. No.:		0		Length: 11230	
Score:		16707.00		Matches: 2915	
Percent Similarity:		89.72%		Conservative: 289	
Best Local Similarity:		81.63%		Mismatches: 361	
Query Match:		83.65%		Indels: 6	
DB:		6		Gaps: 5	
US-09-977-053-4 (1-3571) x AX375163 (1-11230)					
QY	1	MetTTPProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTTPAlaThr	20		
DB	176	ATGTGGTCGGCGCTGGCCCTTTTGTGCTGGGCTCTGGCACTGGTTCGGGCTGGACCAAC	235		
QY	21	PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro	40		
DB	236	TTCCAGCGCGTGGCCCTTGGCTCACTTCAGCTTCGGCTGTYCCCGAGGCTCTCCG	295		
QY	41	GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer	59		
DB	296	GGGGCTCTGGGCAGACTGGCGGTACTCCCGGCTCCAGTGGAGGAGGAGGAGGAGG	355		
QY	60	ArgValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSer	79		
DB	356	AAAGTGGAGCGCTGGCGCGCGTTCGGAGCCGCTGGCGGCGACTGGCGGAGCTCAGC	415		
QY	80	GluArgLeuGluLeuValPheLeuValAspSerSerSerValGlyGluValAlaAsnPhe	99		
DB	416	GGCAGCTCTGGAGCTGCTTCTCTGCTGGAGAGTGTCCAGAGTGGGCCAAACCACTTC	475		
QY	100	ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValProThrAla	119		
DB	476	CTCAACGAGCTCAAGTTTCTGGCGAAGCTGTGTCCGACTTCCCGGTGGTGTCCAGCGCC	535		
QY	120	ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr	139		
DB	536	ACGGGTGGGCGCATCGTCACTTCTCATCCAGAACCAACGTTGGTGGCGCGGTGGATTAC	595		
QY	140	IleSerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAla	159		
DB	596	ATCTCCACGCGCGCGGCACCAACAAAGTGGCGCTGTCTCAGCGCGAGATCCGGCC	655		
QY	160	IleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIle	179		
DB	656	ATCAGCTACCGCGGTGGTGACCTATATACAAAGGCGCTTCCAGCAAGCGCGCAATC	715		
QY	180	LeuLeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSer	199		
DB	716	CTTCGTCACTTAGAGAAACCTCCACCAAGTCATATTTCTATCAGCGAGCTTATTC	775		
QY	200	AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe	219		
DB	776	AATGCGGAGACCCAGACCTATTTCAGCATCGCTTCGGGATTTCCGAGTGAGATCTTC	835		
QY	220	ThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys	239		
DB	836	ACGTTCCGGGATTTGGCAGGGAATATCCGGGAATGAAATGACATGGCTTCCACCCGGAAG	895		
QY	240	GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAla	259		
DB	896	GAAGAACATTGTACTCTCCACAGTTTGAAGAAATTTAGGCTTTAGCTCGCAGGGCG	955		
QY	260	LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr	279		
DB	956	TTGCATGAAGATCTACTCTCTGGAGTTTATCCAGAGGATATGGCCCTCTTCTTAT	1015		
QY	280	LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis	299		
DB	1016	CTCTGTGAGCTGGGAAAGACTGTGTGACAGATGGCCAGCTGCAAAATGTGGGACAC	1075		
QY	300	ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyGlyLeuGlnTyrGlu	319		
DB	1076	ACGGGTCAATTGAAATGATCTGTGAGAGGGCTATTACGGGAAAGGTCTCGAGCATGAG	1135		
QY	320	CysThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSer	339		
DB	1136	TGCACAGCTTGGCCATCAGGAGACATATAAGCGGAGCTTCTCCAGAGGAATCAGCACC	1195		
QY	340	CysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAsp	359		
DB	1196	TGCATCCCATGTCTGACGTAAAGCACACCTCCCACTCCGGAAGCACTTCCCTCGAAGAC	1255		
QY	360	CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysPro	379		
DB	1256	TGGTGTGCCGAGAGGATACAGAGATCTGCCAGACTGTGAGGTGTCCACTGTCTCT	1315		
QY	380	AlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsn	399		
DB	1316	GCCTTGAAGCTCTCTGAAATGTTTTTTTATACAAACACTTGCAGAAATCTACTTCAAT	1375		
QY	400	AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu	419		
DB	1376	GGCGCTGTGGGGTCCGATGTCGCGGCTTTGACCTTGTGGGAAGCAGCATCTTCT	1435		
QY	420	CysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysPro	439		
DB	1436	TGTCACCCATGTGTTTGGTCTGGGACAGAAAGCTTCTGCAGAGTGAGAACGTGCCCC	1495		
QY	440	HisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLys	459		
DB	1496	CACCTCCAGACGCCCAACACAGCCACATCAGCTGTCTCCACTCGCGAAATGTCTACAAC	1555		
QY	460	ThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThr	479		
DB	1556	ACCTGTGTGTGGTGTACCTGCAATGAAGATACAGATTAGAAGCAGCAGCTAGCTTACC	1615		
QY	480	CysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSer	499		
DB	1616	TGTCAGGAAATGCCAGTGGATGGCCAGAGCCCGGTGTGTAGAACGCCATTTGGCC	1675		
QY	500	ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla	519		
DB	1676	ACCTTCCAGAGCCCAAGCGCTCATCTTCTCCACCCAGCTGCGGCAAGCAGCGGCC	1735		
QY	520	LysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLys	539		
DB	1736	AGGCTGGGATGACCTGTGAGTAAAGTCCCGCAGGATACATTTATCCGGGGTGAGA	1795		
QY	540	GluMetLeuArgCysThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCys	559		
DB	1796	GAA---GTGAGATGTCACATCTGGGAAGTGGAGTGGCCAAAGTTCAGACAGCTGTGTGC	1852		
QY	560	LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu	579		
DB	1853	AAAGATGTGGAGGCTCCACAAATCAGCTGTCCAAATGACATTTGAGGCAAGACTGGGGAG	1912		
QY	580	GlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlu	599		
DB	1913	CAGCAGGACTTGCTTAATGCCACTGGCAAGTCCCAAGCTTAAGACCAACTCTGTGAA	1972		
QY	600	LysValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAsp	619		
DB	1973	AAAGTGTGAGTCCAGCTCCACCGCTTTTACCCCACTTACCTCTTCCCAATTTGGAGAC	2032		
QY	620	ValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis	639		
DB	2033	GTGGCCATCACCTACACGCAACCGCACTCATCCGGTAACCAAGCCAGCTGCACTTCTAC	2092		

Db 4250 TCTAAGGATGGTGCACACAGCTTTCAGGTGTCAATGTCTCCAGCAGGCTTTCACAGGGACACAC 4309
Qy 1380 CysGluLeuAasnIleAasnGluCysGlnSerAasnProCysArgAasnGlnAlaThrCysVal 1399
Db 4310 TGTGAACATCAACAGAGTGTCAAGTCCACCCCGTGTAGGAACACAGGCCACCTGTGTG 4369
Qy 1400 AspGluLeuAasnSerTyrSerCysValCysGlnProGlyPheSerGlyGlnArgCysGlu 1419
Db 4370 GATGAACCTAATCATACAGTTGTGAATGTACAGCCAGGATTTTCAGGCCACAGGTGTGAG 4429
Qy 1420 ThrGluGlnSerThrGlyPheAasnLeuAaspPheGluValSerGlyIleTyrGlyTyrVal 1439
Db 4430 ACAGAACACGCTTCCGGTTTAACTCGGATTTGAAGTTTCTGGCATCTACGGGTACGTC 4489
Qy 1440 MetLeuAaspGlyMetLeuProSerLeuHisAlaLeuThrCysThrPheThrPheTyrSer 1459
Db 4490 CTGCTAGATGGAGTGTGCACACCCCTCCATGCGGTAACTGCGCATTCCTGGATGAATCC 4549
Qy 1460 SerAaspMetAasnTyrGlyThrProIleSerTyrAlaValAaspAasnGlySerAasn 1479
Db 4550 TCTGATGTCACTACACGGAGCCCATCTCTATGCACTTGAGGATGACAAAGACAC 4609
Qy 1480 ThrLeuLeuLeuThrAaspTyrAasnGlyTyrValLeuTyrValAasnGlyArgGluIle 1499
Db 4610 ACCTTCTCTCCTGATTAACAAGGCTGGTCTTTTATGTGAATGGAAGGAAGATC 4669
Qy 1500 ThrAasnCysProSerValAasnAaspGlyArgTyrPheHisIleAlaIleThrTrpThrSer 1519
Db 4670 ACCAACTGCCCCCTCCGTAATGATGTCATTTGGCATCATATGCAATCATCATGACACAGT 4729
Qy 1520 AlaAasnGlyIleTrpLysValTyrIleAaspGlyLysLeuSerAaspGlyAlaGlyLeu 1539
Db 4730 ATTGGTGGAGCTGGGGTCTATATAGATGGGAATTAATCTGACCGGTGTACTGGCCTC 4789
Qy 1540 SerValGlyLeuProIleProGlyGlyAlaLeuValLeuGlyGlnGlnAaspLys 1559
Db 4790 TCCATTGGCAAGCCATACCTGGTGGCGTGCATTAGTTCTTGGGCAAGACAAAGACAAA 4849
Qy 1560 LysGlyGluGlyPheSerProIleAgluSerPheValGlySerIleSerGlnLeuAasnLeu 1579
Db 4850 AAAGAGAGGGGTCAACCCGGCTGAGTCTTTGTGGGCTCCATAAGCCAGCTCAACCTC 4909
Qy 1580 TrpAaspTyrValLeuSerProGlnGlnValLysSerLeuAlaThrSerCysProGluGlu 1599
Db 4910 TGGGATATGTCTGTCTCCACAGCAGGTGAAGTCTCTGGCCAGCTCTCTGCCACAGGAA 4969
Qy 1600 LeuSerLysGlyAasnValLeuAlaTrpProAaspPheLeuSerGlyIleValGlyLysVal 1619
Db 4970 CTGAGTGGGGGAAACGTGTAGCATGGCCCGATTTCTGTGGGAATCACGGGGAAGGTG 5029
Qy 1620 LysIleAaspSerLysSerIlePheCysSerAaspCysProArgLeuGlyGlySerValPro 1639
Db 5030 AAGGTTGATTCAGCAGCATGTCTGCTGTGATTTGTCCGTCTTTAGAAGATCCCGTGCCT 5089
Qy 1640 HisLeuAagThrAlaSerGluAaspLeuLysProGlySerLysValAasnLeuPheCysAasp 1659
Db 5090 CACCTGAGACCTGTCATCAGGAATCGAAAGCCAGGCTCCAAAGTCAGTCTGTCTGTGAT 5149
Qy 1660 ProGlyPheGlnLeuValGlyAasnProValGlnTyrCysLeuAasnGlnGlyGlnTrpThr 1679
Db 5150 CCGGGCTTCAGATGGTGGGAATCTGTGTCAGTATTGTCTGAACCAAGGGCAGTGGACA 5209
Qy 1680 GlnProLeuProHisCysGluAargIleSerCysGlyValProProLeuGluAasnGly 1699
Db 5210 CAACCACTCCCACTGTGAAGCATTCGCTGTGGGCTGCTCCCGCCTTGGAGATGGC 5269
Qy 1700 PheHisSerAlaAaspPheTyrAlaGlySerThrValThrTyrGlnCysAasnAasnGly 1719
Db 5270 TTCTACTCAGCGAGACTTCCATGGGGCAGCAGCAGGTGACCTATCAGTGCACCAAGTGGC 5329
Qy 1720 TyrTyrLeuLeuGlyAaspSerArgMetPheCysThrAaspAasnGlySerTrpAasnGlyVal 1739
Db 5330 TACTACCTGCTGGTGAATCCCGAATGTTCTGCRACAGACAAAGGGAGCTGGAAACGGCAT 5389

Qy 1740 SerProSerCysLeuAaspValAaspGluCysAlaValGlySerAaspCysSerGluHisAla 1759
Db 5390 TCACCATCTCTCTCATGTCTGATGTGCGAGTGCAGTCCGCTCGACTCGACTGTAGTACACGCC 5449
Qy 1760 SerCysLeuAasnValAaspGlySerTyrIleCysSerCysValProProTyrThrGlyAasp 1779
Db 5450 TCTGCTCTGAACACCAACGAGTCTCTAGTATGCTCTCTGTAAACCCACCATACACGGAGAT 5509
Qy 1780 GlyLysAasnCysAlaGluProIleLysCysLysAlaProGlyAasnProGluAasnGlyHis 1799
Db 5510 GGGAAACCTGTGCAGAACCTGTAAATGTAAAGCTCCAGAAATCCAGAAATGGCCGC 5569
Qy 1800 SerSerGlyGluIleTyrThrValGlyAlaAlaValThrPheSerCysGlnGluGlyTyr 1819
Db 5570 TCTTCTGGGAGATTTACACCGTGGTACTGCGAGTACATTTCTCTGTGACGAGGGCAC 5629
Qy 1820 GlnLeuMetGlyValThrLysIleThrCysLeuGluSerGlyGluTrpAasnHisLeuIle 1839
Db 5630 GAGCTGGTGGAGTGAAGCACCACCATCAGTGTGTGGAGACTGGCGAGTGGGATCGCTCAGG 5689
Qy 1840 ProTyrCysLysAlaValSerCysGlyLysProAlaIleProGluAasnGlyCysIleGlu 1859
Db 5690 CGGTCTCTGAAAGCCATTTCTGTGTGTCCCACTGTTCCTGAAATGTGTGTGTGAC 5749
Qy 1860 GluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAasnLysGlyTyrThrLeu 1879
Db 5750 GGGTGGCATTCATCATATGCGAGTAAAGTGTGTACAGGTGTGATAAAGATATATCTTTG 5809
Qy 1880 AlaGlyAaspLysGluSerCysLeuAlaAasnSerSerTrpSerHisSerProProVal 1899
Db 5810 TCTGGGATGAAGAGTGAAGTGCCTTGTAGTGTCTCTGGAGTCACTCTCTCTCTGTG 5869
Qy 1900 CysGluProValLysCysSerSerProGluAasnIleAasnAasnGlyLysTyrIleLeuSer 1919
Db 5870 TGGGGCTAGTGAAGTGTTCAGCAGCTCCAGCTGAGGATAAATAACGGCAATATCATCTTAAGT 5929
Qy 1920 GlyLeuThrTyrLeuSerThrAlaSerTyrSerCysAaspThrGlyTyrSerLeuGlnGly 1939
Db 5930 GGGTCACTACTCTTCTTATTTGATGATGCTCTCTGTGAGAACGGATACAGTTTACAGGGC 5989
Qy 1940 ProSerIleIleGluCysThrAlaSerGlyIleTrpAaspArgAlaProProAlaCysHis 1959
Db 5990 CCATCTCTCTGATGATGACAGCTTCGGCAGCTGGGACAGAGCGCCACCTAGCTGTCAA 6049
Qy 1960 LeuValPheCysGlyGluProProAlaIleLysAaspAlaValIleThrGlyAasnAasphe 1979
Db 6050 CTGTCTCTCTGGGAGAGCTCCCAATCGTCAAAAGATGCTGTCTATCTACTGGGAGCACTTC 6109
Qy 1980 ThrPheArgAasnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAasp 1999
Db 6110 ACTTTTGGGACACAGTTCCTTACATGCAAGAGGGCTTACACCTTGTCTGGGCTGAC 6169
Qy 2000 ThrIleGluCysLeuAlaAaspGlyLysTrpSerArgSerAaspGlnGlnCysLeuAlaVal 2019
Db 6170 ACCATCATATGCCAGGCCAACGGCAATGGAATTCAGTAAACCAACAGTGCCTGGCTGTC 6229
Qy 2020 SerCysAaspGluProProIleValAaspHisAaspProGluThrAlaHisArgLeuPhe 2039
Db 6230 TCTGTGACGAGCCGCCCAATGTGGACCAAGCTCTCTCCAGAGACTGTCTCAGAGCTCTTT 6289
Qy 2040 GlyAaspIleAlaPheTyrTyrCysSerAaspGlyTyrSerLeuAlaAaspAasnSerGlnLeu 2059
Db 6290 GGGACACCGGTTTACTACTGTGGGATGGCTACAGCTGTGCTGATATATTCACAGCTC 6349
Qy 2060 LeuCysAasnAlaGlnGlyLysTrpValProProGluGlyGlnAaspMetProArgCysIle 2079
Db 6350 ATCTGCAATGCCAGGGAACTGGGTTCGCCCGCGCGCCAGGCTGTGCGCGCTGCATA 6409
Qy 2080 AlaHisPheCysGluLysProProSerValSerTyrSerIleLeuGlnSerValSerLys 2099
Db 6410 GGTCACTCTGTGAAAAAATCCCCCATCTGTTTCTTACAGCATCTTGGATCTGTGAGCAAA 6469

QY 2100 AlaIysPheAlaIaIaGlySerValValSerPheLysCysMetGluGlyPheValLeuAen 2119
DB 6470 GCAAGTTTGCAGCTCGCTCGGTAGTAGAGCTTCAAGTGCATGGAGGTTTGTGCTGAAC 6529
QY 2120 ThrSerAlaIysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIle 2139
DB 6530 ACTCAGCGAAGATTGAATGCTGAGAGGTGGAGAGTGGAGCCCTTCTCCCTCTCGGTC 6589
QY 2140 GlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTrpAlaSerGly 2159
DB 6590 CAGTGCATCCCGTGCATGCGGAGAGCCTCCCAAGCATCCCAATGGCTACCCGAGTGGG 6649
QY 2160 SerAsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnIysGlyPheTyrIleLys 2179
DB 6650 ACAACATACAGTTTGGGCGGTGGTGGCTACAGCTGCCAAGGATTCATATCAAG 6709
QY 2180 GlyGluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCys 2199
DB 6710 GGGGAGAAGAGACGACGTGTGAGGCCACAGGACAGTGGAGTAAACCCACGCCCTCTGC 6769
QY 2200 HisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThr 2219
DB 6770 CATCTGTCTCTGTAAACGAGCCACCTAAGGTTGAGAACGGCTTCTCGGAGCACACACT 6829
QY 2220 GlyArgIlePheGluSerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGly 2239
DB 6830 GGCAGGACCTTTGAGAGCGAAGCAGGTTCCAGTGCACCCAGGCTATAGGCGCGCGGA 6889
QY 2240 SerProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCys 2259
DB 6890 AGTCTGTGTGTGTTCACAGCCATCGCCACTGGCAGCGACGCCCTCTGTCTCTGC 6949
QY 2260 ValProLeuAspCysGlyLysProProLysGlnAsnGlyPheMetLysGlyGluAsn 2279
DB 6950 ACCCTCTCAACTGTGGGAAACCCCTCCCAATTCAATGGCTTTTGAAGAGAGAAAGC 7009
QY 2280 PheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAsp 2299
DB 7010 TTGAAGTAGGGTCCAAAGGTTCAGTTTGTCTGTAATGAGGAGATGAGCTCGTGTGGAT 7069
QY 2300 SerSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMet 2319
DB 7070 AATCTTGGACTTGCAGAAATCTGGCAAAATGGAGTAAGAACCCGAGTGTGTC 7129
QY 2320 ProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLeuLysGluLeuThr 2339
DB 7130 CCCACCAAGTGTGCAGAGCTCTCTCTTTAGAAAACCAAGCTCGTATTTGAAGGAATTAGCT 7189
QY 2340 ThrGluValGlyValValThrPheSerCysLysGluGlyHisValLeuGlnGlyProSer 2359
DB 7190 TCCGAGGTAGGAGTATGATGACCAATTTCTGTGAAGAGGGGATGCTTGCAGGCCCTCT 7249
QY 2360 ValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLysIleVal 2379
DB 7250 GTCTCAGTGTCTGCCATCCGGCAATGGAATGTTCTCTTCTATTTTGAAGATGGTC 7309
QY 2380 LeuCysThrProProLeuLeuSerPheGlyValProIleProSerSerAlaLeuHis 2399
DB 7310 CTTTGTCCCTCGCTCTCCCTTGATTTCTTTCGGGCTGCCCTCGCTTTCGGTGTCTTCAT 7369
QY 2400 PheGlySerThrValLysTyrSerCysValGlyGlyPhePheLeuArgGlyAsnSerThr 2419
DB 7370 TTGGCAGTACTGTCAAGTATCTGTGTGTCGACGGGTTTTTCTTAAGAGGAGTCCNACC 7429
QY 2420 ThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValProValGlu 2439
DB 7430 ATCTCTCGCAGGCTGATAGCAGCTGAGTTCTCCATTTGCCGAATGCGTTCGGTAGAA 7489
QY 2440 CysProGlnProGluGluLeuProAsnGlyIleIleAspValGlnGlyLeuAlaTyrLeu 2459
DB 7490 TGTCCCAACCTCAGGAGATCCTCAACGGGTATCATCCAGTACAGGGCTTGCCTATCTC 7549
QY 2460 SerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeu 2479

DB 7550 AGCACCAAGCTCTACACCTGCAAGCCAGGCTTTGAGTTAGTGGCAATGCTACCAACCTC 7609
QY 2480 CysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCysLeu 2499
DB 7610 TGTGGGMAAATGGCCAGTGGCTCGGAGGAAACCAATGTGCAAAACCATTTGAATGCCCA 7669
QY 2500 LysProLysGluLeuLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThr 2519
DB 7670 GAGCCCAAGAGATTTTAAATGGCCAATCTTTCCTGAGCTTTCAGTATGGCAAAACC 7729
QY 2520 ValThrTyrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeu 2539
DB 7730 ATCACAATCTTTGTGACCGGGCTTCGGCTCGAAGGTCCCAATCCCTGACCTGTGTTA 7789
QY 2540 GluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAspSerPro 2559
DB 7790 GAGACAGGTGACGTGGATATGATCCCTCTTGTGATGATCCACTGCAGTGACCCCA 7849
QY 2560 GlnProIleGluAsnGlyPheValGluGlyAlaAspTyrSerTyrGlyAlaIleIleLeu 2579
DB 7850 CAGCCCATTTGAAATGGTTTCTGTAAGAGGTGGGATTTACAGATACGGTGCCTGATCATC 7909
QY 2580 TyrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGlySer 2599
DB 7910 TATAGTGTCTTCTCTGGTTTCAGGTGCTTGGTCAATGTCAGAGCTGTGAGAGTGC 7969
QY 2600 GlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuProProHisIle 2619
DB 7970 GGTGTGTCAAGCTCCAGCCCAACTGTGTACCCATAGACTCGGTCTCTCTCTCCACATA 8029
QY 2620 AspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGlnGluAspAsp 2639
DB 8030 GACTTTGGTGACTGTACTAAAGTCAGAGATGGCCAGGAGCATTTTGTATCAGAGATGAC 8089
QY 2640 MetMetGluValProTyrValThrProHisProProTyrHisLeuGlyAlaValAlaLys 2659
DB 8090 ATGATGGAGTCCCATATCTGGCT-----CACCTCAACATTTGGAAGCAACAGCTAAG 8143
QY 2660 ThrTrpGluAsnThrLysGluSerProAlaThrHisSerSerAsnPheLeuTyrGlyThr 2679
DB 8144 GCCTTGGAAAATCAAAAGGAGTGCCTGCTGCATCATGTCATCCCACTCTCTCTATGGCAG 8203
QY 2680 MetValSerTyrThrCysAsnProGlyTyrGluLeuLeuGlyAsnProValLeuLysCys 2699
DB 8204 ATGTTTCTTACAGCTGCGAGCTGTGTATGAACTGCTGGAAATCCCTGCTGCTGATCTGC 8263
QY 2700 GlnGluAspGlyThrTrpAsnGlySerAlaProSerCysIleSerIleGluCysAspLeu 2719
DB 8264 CAGGAAGTGGTACGTGGAAATGTTACCGCACCTCTTTCATTTCCATTTGAATGTGATTTG 8323
QY 2720 ProThrAlaProGluAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaVal 2739
DB 8324 CTTGTTGCTCCCGAAAATGGCTTTTACATTTTACACAGACGACTATGGGCGAGTGTGCA 8383
QY 2740 GlnTyrSerCysLysProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGlu 2759
DB 8384 CAATATAGCTGCAGCGGGGACATTTCTAAGAGGCTCCCACTTAAGACTCTGTCTGTCAG 8443
QY 2760 AspArgLysTrpSerGlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsn 2779
DB 8444 AATAAGCAGTGGAGTGGCACTGTTCCAGCTGTGTAAGCCATCTCATGCAAGTAAAGCAAC 8503
QY 2780 ProValMetAsnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyr 2799
DB 8504 CCACCTCGGAATGGATCCATCAAGAGGAGATGACTACTCTACCTGGGTGTGTGTACTAC 8563
QY 2800 GluCysAspProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLys 2819
DB 8564 GAGTGTGACTCTGGCTATATCTCAATGGCTCTAAGAAAGAGGATGCTCCAAAGAAATAGA 8623
QY 2820 AsnTrpAspGluAspGluProIleCysIleProValAspCysSerSerProProValSer 2839

Db	8624	GATTGGATGGCGATGAGCCATGTGTATTCTCTGTAGACTGTGGCTCACCCCGAGTCCC	8683
Qy	2840	AlaAenGlyClnValArgGlyAspGlnTyrThrPheGlnLysGluIleGlnTyrThrCys	2859
Db	8684	ACCAATGGCCGAGTGAAGGAGGAAGATACACATTCCTCAAAAGGAGATACATCTCTGC	8743
Qy	2860	AsnGluGlyPheLeuLeuGluGlyAlaAqSerArgValCysLeuAlaAenGlySerTyr	2879
Db	8744	CGTGAAGGGTTCATCTGGNAGGAGCCAGGAGTCGTATCTGTCTTACCAATGGAGTTGG	8803
Qy	2880	SerGlyAlaThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAen	2899
Db	8804	AGTGGTGCCTCCAGCTGCATGCTGTAGATGCTCTGCCCCACCAACAGGTGCCAAT	8863
Qy	2900	GlyValThrGluGlyLeuAspTyrGlyPheMetLysGluValThrPheHisCysHisGlu	2919
Db	8864	GGGGTGGCAGATGGCTAGACTATGGGTTCAAGAAGGAGTAGTCCACTGTCTAGAG	8923
Qy	2920	GlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGlyAsnTyrAspAla	2939
Db	8924	GGCTATGTCTGCAGGGGGCTCCAGACTCCTCTGTCTAGTCCCAATGGGACTTGGGATGCA	8983
Qy	2940	GluIleProLeuCysLysProValAsnCysGlyProProGluAspLeuAlaHisGlyPhe	2959
Db	8984	GAAGTCCCTCTGTAAACAGCTACCTGTGGTCTCTCTGCCGACCTTCCCAGGGCTTC	9043
Qy	2960	ProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrClnCysPheProGlyTyr	2979
Db	9044	CCTAATGGCTTTCTTTTATCATATGGGGGCCACATACAGTATCAGTGTCTTACTGGTTAT	9103
Qy	2980	LysLeuHisGlyAsnSerSerArgAqCysLeuSerAsnGlySerTyrPheGlySerSer	2999
Db	9104	AGCTTTCATGGAACCCATCAAGAGATGCTTCCCAATGGTCTCTGGAGCGGCGACTCG	9163
Qy	3000	ProSerCysLeuProCysArgCysSerThrProValIleGlnTyrGlyThrValAsnGly	3019
Db	9164	CCATCTGCTCTCTGAGGTGTTCCACACCCATCATCAACAGAGCAACCATCAACGCA	9223
Qy	3020	ThrAspAspCysGlyLysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeu	3039
Db	9224	ACTGATTTGGGATGTGGAAGAGCGTCCAGATTTGAGTGTCTCAAGGCTTCAAGCTGCTT	9283
Qy	3040	GlyLeuSerGluIleThrCysGluAlaAspGlyGlnTyrSerSerGlyPheProHisCys	3059
Db	9284	GGACTTCTGAATCATCTGTGATGCCAATGGCTCTGAC---GTCCACCTGTGT	9340
Qy	3060	GluHisThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSer	3079
Db	9341	GAGCAGCTCAGTGGCGGCTCTCCCAACCATACCCAAACGCAATTTGCTTGGAGGGCAGC	9400
Qy	3080	SerTyrLysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySer	3099
Db	9401	CTTTCCGAGGACAATGTGTAACTTACAGCTGCAGACCTGGCTACACCATGCAAGTAGT	9460
Qy	3100	SerAspLeuIleCysThrGluLysGlyValTyrPheSerGlnProTyrProValCysGluPro	3119
Db	9461	TCGATCTGATTTGTACGGAAGGCAATATGAGGCGAGCTTACCAACCTGTGAACCC	9520
Qy	3120	LeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThr	3139
Db	9521	CTGTCTGTGGACCCCACTGTAGCCAAATGACAGTGCAGTGCAGCAACAGGAGGCTCATACC	9580
Qy	3140	TyrGluSerGluValLysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAsp	3159
Db	9581	TATGAAGCAAGTGAATCTCAGGTCTCGAAGGGTATGTGATGATTCGGATACAGAT	9640
Qy	3160	ThrPheThrCysGlnLysAspGlyArgTyrPhePheProGluArgIleSerCysSerProLys	3179
Db	9641	ACATTCACCTGCCAGCAAGATGGCCATATGGTCCCTGGAAGAATCACCTGCAGTCTTAA	9700
Qy	3180	LysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerVal	3199
Db	9701	AAATGCCCTGTGCATCCCAACATGACACGATACGTTTTCACGGAGATGACTTCCAGGTG	9760
Qy	3200	AsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAsnIleSer	3219
Db	9761	AACAGACAGATTTCTGTGTCTGTGTGATGTGAGAGGGTTTACCCAGGAGGATGACTGGTCA	9820
Qy	3220	ValCysGlnLeuAspGlyThrTyrGluProProPheSerAspGluSerCysSerProVal	3239
Db	9821	ACATGCCAGCCGCGATACATGGGAGCCACCATTTTCTGATGATCTCTGTATCCAGTT	9880
Qy	3240	SerCysGlyLysProGluSerProGluHisGlyPheValValGlySerLysTyrThrPhe	3259
Db	9881	GTTTGTGGGCATCTCTAAAGCCAGCCCATGGCTCCGTGGTGTGGCATAAAGACACAGCTTT	9940
Qy	3260	GluSerThrIleIleTyrGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg	3279
Db	9941	GGAGACCATTTGTTTACAGTGTGTACCTGGCTACAAATAGAGGGGACAGGGAACGA	10000
Qy	3280	ValCysGlnGluAsnArgGlnTyrSerGlyGlyValAlaIleCysLysGlyThrArgCys	3299
Db	10001	ATCTGCCAGGAGAACAGACAGTGGAGTGGAGAGTGGCAGTGTGCAGAGAGAACAGATGT	10060
Qy	3300	GluThrProLeuGluPheLeuAsnGlyLysAlaAspIleGluAsnArgThrGlyPro	3319
Db	10061	GAGACTCCAGCTGAGTTTCCCAATGGGAAGGCTGTCTTGGAAAACACACATCTGGACCC	10120
Qy	3320	AsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCys	3339
Db	10121	AGCTTCTGTCTTCTGTACAGAGGCTACACCTGGNAGGGTCCCGGAGGCACACTGC	10180
Qy	3340	ThrGluAsnGlyThrTyrPheSerHisProValProLeuCysLysProAsnProCysProVal	3359
Db	10181	ACTGCAATGGAACTCTGGAATCACCTGCTCTGCAAAACCAATCCATGCTCTGTC	10240
Qy	3360	ProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGluAsn	3379
Db	10241	CCTTTTGTGATCTCTGAGAACGCGCTCTTCTGAAAAGAGTTTATGTGCGACAGAT	10300
Qy	3380	ValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsn	3399
Db	10301	GTATCTATCAGTGCAGGAGGCTCTCTGCTCAAGAGCAATGTGTGTCTCATCAGTGCAGC	10360
Qy	3400	ProAspGluThrTyrPheThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProPro	3419
Db	10361	CCTGACGACATGAGCAGCACCAATGCCAGATGTGAAAAAATCTCTGTGTGCTCTCCA	10420
Qy	3420	AlaHisValGluAsnAlaIleAlaGlyValHisTyrGlnTyrGlyAspMetIleThr	3439
Db	10421	AGTCACGTAGAGAAATGCAATTTGCTCAGAGAGTGTATTACCAAGTATGGGGACATGATCACC	10480
Qy	3440	TyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsn	3459
Db	10481	TACTCTGTTCAGTGGCTACATGCTAGAGGTTCTCTCGGAGTGTTCCTAGAGAAAT	10540
Qy	3460	GlyThrTyrThrSerProPheCysArgAlaValCysArgPheProCysGlnAsnGly	3479
Db	10541	CGAACATGACACCATCTCTCTGTTTCAGAGCTGTCTGTCCGTTCCCATGCGAATGGA	10600
Qy	3480	GlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCys	3499
Db	10601	GGTGTCTGTCAACGTCCAAATGCTCTCATGCCCCAGAGCGCTGGAGTGGACGCTCTGT	10660
Qy	3500	GluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaIleProTyrGln	3519
Db	10661	GAAGACCAATATGATATCTCTCTCTTGTGATGGTGGGCGCTGTGTGGCCCTTATCAG	10720
Qy	3520	CysAspCysProProGlyThrThrGlySerArgCysHisThrAlaValCysGlnSerPro	3539
Db	10721	TGTGACTGCCACAGGCTGGAGTGGTCCCGCTGTCTATACAGCTACTTGTCTAGTCCCCC	10780
Qy	3540	CysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeuSerSerTyrThr	3559
Db	10781	TGCTTAAATGGCGGGAAATGTCATAAGCAACCAACCGATGCCATTTGTCTCTGAGCTGACCA	10840

2310 TrpAsnLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeu 2329
2330 GluAsnGlnLeuValLeuLysGlnLeuThrThrGluValGlyValValThrPheSerCys 2349
2340 GAAACCACTAGTATTAAAGGAGTTGACCAAGGAGTGGAGTTGTGACATTTTCTGT 1501
2350 LysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrp 2369
2370 AsnAspSerPheProValCysLysLysLysValLeuLysCysThrProProLeuLysSerPhe 2389
2390 GlyValProLeuProSerSerAlaLeuHisPheGlySerThrValLysTrpSerCysVal 2409
2410 GlyLysPheLeuArgLysAsnSerThrThrLeuCysGlnProAspGlyThrTrpSer 2429
2430 SerProLeuProGluCysValProValGlnCysProGlnProGluGluLeuProAsnGly 2449
2450 IleLeuAspValGlnGlyLeuAlaValLeuSerThrAlaLeuThrThrCysLysProGly 2469
2470 PheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGly 2489
2490 LysProThrCysLysAlaIleGluCysLeuLysProLysGluLeuAsnGlyLysPhe 2509
2510 SerThrThrAspLeuHisThrGlyGlnThrValThrThrSerCysAsnArgGlyPheArg 2529
2530 LeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaPro 2549
2550 SerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGly 2569
2570 AlaAspThrSerThrGlyAlaIleIleThrSerCysPheProGlyPheGlnValAla 2589
2590 GlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerSerIleProThrCysMet 2609
2610 ProIleAspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAsp 2629
2630 AspGlnGlyThrPheGluGlnGluAspAspMetMetGluValProThrValThrProHis 2649
2650 ProProThrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAla 2669
2670 ThrHisSerSerAsnPheLeuThrGlyThrMetValSerThrThrCysAsnProGlyThr 2689
2690 GluLeuLeuGlyValAsnProValLeuLysCysGlnGluAspGlyThrTrpAsnGlySerAla 2709
2710 ProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArg 2729
2730 PheThrGluThrSerMetGlySerAlaValGlnThrSerCysLysProGlyHisIleLeu 2749
2750 AlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaLysProArg 2769
2770 CysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySer 2789
2790 AsnThrThrThrLeuSerThrLeuThrValGlnCysAspProGlyThrValLeuAsnGly 2809
2810 ThrGluArgArgThrCysGlnAspLysAsnTrpAspGluAspGluProIleCysIle 2829
2830 ProValAspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluThr 2849
2850 ThrPheGlnLysGluLeuGlyThrCysAsnGluGlyPheLeuLeuGluGlyAlaArg 2869
3002 ACATTCGAGGAGGAGTGAATGACATTCGATGAAGGGTCTTGCTGGAGGAGCCAGG 3061
3062 AGTCGGGTTTGTCTGCAATGGAGTGGAGTGGAGCCACTCCGCACTGTGTGCTGTGTC 3121
3080 ArgCysAlaThrProGlnLeuAlaAsnGlyValThrGluGlyLeuAspThrGlyPhe 2909
3122 AGATGTGCCACCCCGCCACACTGGCCCAATGGGGTGACGGGAAGGCTTGGACTATGGCTTC 3101
2910 MetLysGluValThrPheHisCysHisGluGlyThrIleLeuHisGlyAlaProLysLeu 2929
3182 ATGAGGAGGAGTAAACATTCCTCCTCAGAGGGCTACATCTTGCCAGTGTCTCCAAACTC 3241
2930 ThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCys 2949
3242 ACCTGTCACTGAGTGGCACTGGGATGAGAGATTCCTCTCTGTAAACCACTCACTGT 3301
2950 GlyProGluAsnLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGly 2969
3302 GGACCTCTCTGAGAGATCTTCCCATGTTTCCCTTAATGGTTTTTCTTATTCATGGGGGC 3361
2970 HisIleGlnThrGlnCysPheProGlyThrLysLeuHisGlyAsnSerSerArgArgCys 2989
3362 CATATACATGATCAGTGCTTCTCTGGTTATTAAGTCCATCGAATTCATCAGAAGGTGC 3421
2990 LeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThr 3009
3422 CTCTCCATATGGCTCTCTGGAGTGGAGTCACTTCTCTGCTGCTCTTTCAGATGTTCACA 3481
3010 ProValIleGluThrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaLysArg 3029
3482 CCAGTAAATGGAATATGGAATCTGCAATGGAGACAGATTTTGACTGTGGAAAGGAGCCGG 3541
3030 IleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAsp 3049

2670 ThrHisSerSerAsnPheLeuThrGlyThrMetValSerThrThrCysAsnProGlyThr 2689
2462 ACATTCATCAAACTTTCTGTATGTATACCATGGTTTTCATACACCTGTATATCCAGGATAT 2521
2690 GluLeuLeuGlyValAsnProValLeuLysCysGlnGluAspGlyThrTrpAsnGlySerAla 2709
2522 GAACTTCTGGGAGACCTGTGTGTGATCTGCCAGAGAGATGGAACCTAGGATGGCAGTGCA 2581
2710 ProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArg 2729
2582 CCATCTCGCATTTCAATTCGAATGTGACTTGCCTTACTCTCTGAAATATGCTTTTTCGCT 2641
2730 PheThrGluThrSerMetGlySerAlaValGlnThrSerCysLysProGlyHisIleLeu 2749
2642 TTTACAGAGACTAGCATGGAAGTGTCTGTGAGTATAGCTGTAAACCTGAGACATTTCTA 2701
2750 AlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaLysProArg 2769
2702 GTGGGCTCTGGCTTAAGGCTTTGTCTAGAGAAATAGAAAGTGGAGTGGTGCCTCCCAAGC 2761
2770 CysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySer 2789
2762 TGTGAAGCCATTTTCATGCAAAAGCCAAATCCAGTCAATGAATGGATTCATCAAGGAAGC 2821
2790 AsnThrThrThrLeuSerThrLeuThrValGlnCysAspProGlyThrValLeuAsnGly 2809
2822 AACTACACATACCTGAGCAGCTGTACTATGATGATGTGACCCCGATATGTGTAATGGC 2881
2810 ThrGluArgArgThrCysGlnAspLysAsnTrpAspGluAspGluProIleCysIle 2829
2882 ACTGAGAGGAGAACATGCCAGGATGACAAAACCTGGGATGAGGATGAGCCCATTTGCATT 2941
2830 ProValAspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluThr 2849
2942 CTGTGTGAGCCGAGTTTCACTCCAGTCTCAGCCCAATGGCCAGGTGAGAGAGACGAGTAC 3001
2850 ThrPheGlnLysGluLeuGlyThrCysAsnGluGlyPheLeuLeuGluGlyAlaArg 2869
3002 ACATTCGAGGAGGAGTGAATGACATTCGATGAAGGGTCTTGCTGGAGGAGCCAGG 3061
2870 SerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValProVal 2889
3062 AGTCGGGTTTGTCTGCAATGGAGTGGAGTGGAGCCACTCCGCACTGTGTGCTGTGTC 3121
2890 ArgCysAlaThrProGlnLeuAlaAsnGlyValThrGluGlyLeuAspThrGlyPhe 2909
3122 AGATGTGCCACCCCGCCACACTGGCCCAATGGGGTGACGGGAAGGCTTGGACTATGGCTTC 3101
2910 MetLysGluValThrPheHisCysHisGluGlyThrIleLeuHisGlyAlaProLysLeu 2929
3182 ATGAGGAGGAGTAAACATTCCTCCTCAGAGGGCTACATCTTGCCAGTGTCTCCAAACTC 3241
2930 ThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCys 2949
3242 ACCTGTCACTGAGTGGCACTGGGATGAGAGATTCCTCTCTGTAAACCACTCACTGT 3301
2950 GlyProGluAsnLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGly 2969
3302 GGACCTCTCTGAGAGATCTTCCCATGTTTCCCTTAATGGTTTTTCTTATTCATGGGGGC 3361
2970 HisIleGlnThrGlnCysPheProGlyThrLysLeuHisGlyAsnSerSerArgArgCys 2989
3362 CATATACATGATCAGTGCTTCTCTGGTTATTAAGTCCATCGAATTCATCAGAAGGTGC 3421
2990 LeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThr 3009
3422 CTCTCCATATGGCTCTCTGGAGTGGAGTCACTTCTCTGCTGCTCTTTCAGATGTTCACA 3481
3010 ProValIleGluThrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaLysArg 3029
3482 CCAGTAAATGGAATATGGAATCTGCAATGGAGACAGATTTTGACTGTGGAAAGGAGCCGG 3541
3030 IleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAsp 3049


```

3542 ATTCAAGGCTTCAAGGCTTCAAGGCTTCTAGGACTTTCTGAAATCACCCTGTGAAGCCGAT 3601
3050 GYGLNTPSRSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMet 3069
3602 GGCAGTGGAGCTCTGGGTTCCGCCACTGTGAACACACTTCTTGTTGGTTCTCTTCCAATG 3661
3070 IleProAsnAlaPheIleSerGluThrSerSerThrPlyGluAsnValIleThrTyrSer 3089
3662 ATACCAAAATCGTTTCATCAGTGAACAGCAGCTCTTGGAAAGGAAAATGTGATAACTTACAGC 3721
3090 CysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyVal 3109
3722 TGCAGTCTGGATATGTATACAGGAGGAGTTCAGATCTGATTTGTACAGAGAAAGGGGTA 3781
3110 TrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAla 3129
3782 TGGAGCCAGCTTATCCAGTCTGTGAGCCCTGTCTGTGGTCCCAACCGTCTGTGCC 3841
3130 AsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLeu 3149
3842 AATGCAAGTGGCAACTGGAGAGGACACACACTATGAAGTGAAGTGAAGTGAAGTGAAGT 3901
3150 GluGlyTyrThrMetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArgTyr 3169
3902 GAAGGTTTATACGATGAGATACAGATACAGATACATTCACCTGTGAGAGAGATGTCGCTGG 3961
3170 PheProGluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleThrHis 3189
3962 TTCCTGAGAGATCTCTGAGTCTTAAATAATGTCCTCTCCGGAAACATACATACAT 4021
3190 IleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGlu 3209
4022 ATACTTGTTCATGGGACGATTTCAAGTGTGAATAGGCAAGTTCTGTGTGTCATGTGCAGAA 4081
3210 GlyTyrThrPheGluGlyValAlaSerValCysGlnLeuAspGlyThrTyrGluPro 3229
4082 GGGTATACCTTTGAGGGAGTTTACATATCATCATCATCATCATCATCATCATCATCATCAT 4141
3230 ProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGluHis 3249
4142 CCATCTCCGATGATCTTCAGTCCAGTTCTTGTGGAAACCTGAAAGTCCAGAACAT 4201
3250 GlyPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluPro 3269
4202 GGATTTGTGTTGGCAGTAAATACACTTTCGAAAGCAGCATTTATTTATCATGTTGAGCCT 4261
3270 GlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTyrSerGly 3289
4262 GGTATGAACTAGAGGGGAAACAGGGAACGTGTCTGCCAGGAGAACACAGATGGAGTGA 4321
3290 GlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLys 3309
4322 GGGGTGGCATATGCAAGAGACACAGGTGTGAACCTCACTTGAATTTCTCAATGGGAAA 4381
3310 AlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyr 3329
4382 GCTGACATTTGAAACACAGGACGACTGGACCCCAAGCTGTGTATATTTCTGCAACAGAGGCTAC 4441
3330 SerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTyrSerHisProVal 3349
4442 AGTCTTGAAGGGCCATCTGAGGACACACTGCACAGAAAATGGAACCTGGAGCCACCCAGTC 4501
3350 ProLeuCysLysProAsnProCysProValProPheValIleProGluAsnAlaLeuLeu 3369
4502 CCTCTCTGCAACCAATCCATGCGCTGTCTCTTGTGATTCGAGATCCGAGATGCTCTGCTG 4561
3370 SerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeu 3389
4562 TCTGAAAGAGGATTTATGTGTATGATCAGATGTGTCTCATCAATGTAGGGAAGGTTTCTG 4621
3390 LeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTyrThrGlnThrSerAla 3409

```

```

4622 CTGCGGGCCAGCGCATCATTTACTACACCCCGACGAGCGTGGACACACAGCAAGCGCC 4681
3410 LysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGly 3429
4682 AAATGTGAAAAAATCTCATGTGGTCCACCGCTCAGGTAGAAAAATGCAATTTGCTCGAGC 4741
3430 ValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGlu 3449
4742 GTACATTTATCAATATGAGACATGATCACCTACTCATGTATACAGTGGGATACATGTTGGAG 4801
3450 GlyPheLeuArgSerValCysLeuGluAsnGlyThrTyrThrSerProProIleCysArg 3469
4802 GGTTCCTGAGAGTGTGTTTGTAGAAAATGGAACATGGACATCACCTCTCTTATTTGCAGA 4861
3470 AlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSer 3489
4862 GCTGTCTGTGATTTCCATGTCAAGATGGGGCATCTGCCAACGCCCAAAATGCTTGTTC 4921
3490 CysProGluGlyTyrMetGlyArgLeuGluGluProIleCysIleLeuProCysLeu 3509
4922 TGTCCAGAGGGCTGGATGGGGCGCTCTGTGAAGAACCAATCTGCATTTCTCCCTGTCTG 4981
3510 AsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProGlyTyrThrGlySer 3529
4982 AACGAGTGTGCTGTGTGGCCCTTACAGTGTGACTGCCGCTGCTGGACGGGTCT 5041
3530 ArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyLysCysValArgPro 3549
5042 CGCTGTCTATACAGCTGTTCGCCAGTCTCCCTGCTTAAATGGTGAATAATGTGTAAGACA 5101
3550 AsnArgCysHisCysLeuSerSerTyrThrGlyHisAsnCysSerArgLysArgThr 3569
5102 AACCGATGTCACTGTCTTTCTTCTGACAGGACATAACTGTTTCAGGAAAGGAGGACT 5161
3570 GlyPhe 3571
5162 GGGTTT 5167

RESULT 9
AK075235
LOCUS Homo sapiens cDNA FLJ90754 fis, clone PLACE4000354, weakly similar
DEFINITION TO COMPLEMENT RECEPTOR TYPE 1 PRECURSOR.
ACCESSION AK075235
VERSION AK075235.1 GI:22761191
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
Sugiyama, T., Suzuki, Y., Nagai, K., Sugano, S., Ishii, S.,
Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
Kojima, S., Nagahara, K., Masuhara, Y., Ono, T., Okano, K., Yoshikawa, Y.,
Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
Ninomiya, K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 5421)
Isogai, T. and Otsuki, T.
Direct Submission
TITLE NEDO human cDNA sequencing project
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

```



```
source 1..5421
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE400354"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="cloning vector: pME18SFL3"
677..5170
/note="unnamed protein product"

CDS
/codon_start=1
/protein_id="BAC11489.1"
/db_xref="GI:22761192"
/translation="MPRLCAHCEKPPSPVSVLSVSKAKPAAGSVSVFKNEGFVL
NTSAKIECRGQWNPSPMSIQCIPIVRCGEPSPIMNGVAGSSVNSFGAMVAYSCNKG
YIKGKKSTCEASQWSPPIPTCHPVSCBEPKVENGLEHTGTIRFSEVRYQCNPG
YKPGSPVFPVCHANEHMSIESPLMWPLDCKGPPPTQNGFMKNGFENFVSKVQFPCNE
GYELVGDSPVTCOKSGKNKSNPKMCAKCEPPLLELNOLVLELTATTEVGVVTPSK
EGHVLQSPVLKCLPSQWNPSPVCKIVLCTPPLISFGVPIPSALHRSSTVYSC
VGGFLRGNTSLQPDGWNSSPLECPVCEPQBEPIPNGLIIVRGLAYLSTALYTC
RPGFLVGNNTTLCEHNGHMLGKPTCKAIECLKPKHILNGKPSYTDLYGQVTVYSC
NRGFLRGSALTCLLETGMDVDAPSCNAIHCDSPQIENGFEVADYSYGAIIYSC
PFGFVAGHAMOTCRSGWSSSIPTCMPIDCGPLHIDFGDCTKLRDQGYEQBQDDM
MEVPVYTPHPYHCAVAKTWENTKESPATHSNFLYGTWVSYTCNPGYELNGVLI
COEDTRNGSAPSCISIEDCLTPAENGFLRTETSMGSAVOYSCPKGHILNGSLRL
CLENRWGASPCRAISCKENPVNMGSIKGNNTYLTSLIYECDPGVLNGSTERRT
CQDDKWDDEDPICIPVDRSSPVSAQGVQVDRYTPQKIBEYTCNEGFLLEGARSV
CLANGSGGATDVPVRCATPGLANGTEGLDYGFMEVTFHCEGYIILHAGPKLT
CQDGNWDAIEPLCKFPVNGCPEDLAHGPNGFSPITGHGHIQYQCPFGYKLHNSRR
CLNSGWSGSSPSCLPFCSTPIVIEGTVNGTDPOGKAARIQCFKFKLGLSITIC
RADQWSSGPFCEHTSCGLSPMINAPISSETSMKENVITYCSGSGVYVIGSSDLIC
TEKGVNSQPIYCEFLSCSPSVANAVATGNAHTYESVVKLRCLLEGTYMDTDTFTT
CQDGRWPFERISCPKCPLENITLILVHGDDFSVKNQSVSCABGTFPEGVNISV
COLDGTWPFPSDESQSPVSCGSPGEPHGVVGYKYTFESTIYQCFEYLEGRE
RVCQNRWGSQVPAICKECTCEQLEFLANGKADINRTTGNVAVSCNRGYLEGPE
ACTENGTVSHVPVLCNPNPCVPFVIPENALLSKPEFVVDQNSIKCRBGLLQGHG
LITVNPDETWTCTSAKCKISGPPAHVENAARGVHYQYGDPMITVYSCYGYMLGFL
RSVCLENGWTSPPIICRAVCRPCQNGGICQRPNACSPEGWMLRCLBERPICILPLN
GRCVAPYQDCDPPGWTGSRCHTAVCQSPCLNGKCKVRPNRCHCLSSWTGHNCRRRR
TGF"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 5421
Score: 9400.00 Matches: 1676
Percent Similarity: 97.56% Conservative: 4
Best Local Similarity: 97.33% Mismatches: 11
Query Match: 47.06% Indels: 32
DB: 9 Gaps: 1

US-09-977-053-4 (1-3571) x AK075235 (1-5421)

QY 1881 GlyAspLysGluSerSerCysLeuAlaAsnSerSerTrpSerHisSerProValCys 1900
DB 3 GGTGATAAGAAATCATCTGCTCTGCTAACAGTTCTTGGAGTCAATTCCTCTCTGTTGT 62
QY 1901 GluProValLysCysSerSerProGluAlaAsnIleAsnAsnGlyLysTrpIleLeuSerGly 1920
DB 63 GAACCAAGTGAAGTGTCTAGTCGGAATAATAAATAATGAAATAATATATTGAGTGGG 122
QY 1921 LeuThrTrpLeuSerThrAlaSerTrpSerCysAspThrGlyTrpSer 1936
DB 123 CTTACTACCTTTCTACTGCAATCATATTATGCGATACAGGATACAG-ATGCGGTTTCG 181
QY 1936 ----- 1936
DB 182 CCATTTTGTCCAGACTAGTCTCGAACTCTTGAGCTCAAGCAATCCACCTGCTCAACCTC 241
QY 1937 -----LeuGlnGlyProSerIleIleGluCysThrAlaSerGly 1949
DB 242 CCAAAAGTGTGGAATTTACAGCTTTACAGGCGCTTCCATTATTGAATGACAGCGCTTCTGCG 301
QY 1950 IleTrpAspArgAlaProProAlaCysHisLeuValPheCysGlyGluProProAlaIle 1969
|||||:::|||||
302 ATCTGGGCAAAAGCGCCACCTGCTCCTGTCACCTCGTCTTCTGTGGAGAACCACTGCGCATC 361
1970 LysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsnThrValThrTrpThrCys 1989
362 AAAGATGCTGTCATTAACGGGGAATAACTTCACTTTTCAGGACACACCGCTCACTTACACTTGC 421
1990 LysGluGlyTrpThrIleuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrp 2009
422 AAAGAAGGCTATACTCTTGTCTGCTCTTGACACCATTTGAATGCTGCGCCGACGCAAGTGG 481
2010 SerArgSerAspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHis 2029
482 AGTAGAGTACACAGAGTGCCTGCTGCTCTGTGATGAGCACCCCACTTTGTGGACCA 541
2030 AlaSerProGluThrAlaHisArgLeuPheGlyAspIleAlaPheTrpTrpCysSerAsp 2049
542 GCCTCTCCAGAGTCTGCCCATCGGCTCTTTGGAGACATTTGCATTTCTACTACTGCTCTGAT 601
2050 GlyTrpSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValPro 2069
602 GGTACAGCTTAGCAGACAATTCACAGCTTCTTGCATGCCCCAGGCAAGTGGGTACCC 661
2070 ProGluGlyGlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerVal 2089
662 CCGAAGGTCACAGACATGCCCCGTTGTATAGTCACTTTCTGTGAANAACCTCCATCGGTT 721
2090 SerTrpSerIleLeuGluSerValSerTrpAlaLysPheAlaAlaGlySerValValSer 2109
722 TCCTATGATCATCTTGGAAATCTGTGAGCAAGCAAAATTTGCAGTGGCTCAGTTGTGAGC 781
2110 PheLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGly 2129
782 TTAAATGATCATGGAAGGCTTTGTACTGAACACCTCGACAAAGATTAATGTATGAGAGGT 841
2130 GlyGlnTrpAsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluPro 2149
842 GGGAGTGAACACCTTCCCCCATGTCATCCAGTGCATCCCTGTGCGGTGTGAGAGCCA 901
2150 ProSerIleMetAsnGlyTrpAlaSerGlySerAsnTrpSerPheGlyAlaMetValAla 2169
902 CCAAGCATCATGAATGGCTATGCAAGTGCATCAACTACAGTTTGTGAGCATGGTGGCT 961
2170 TrpSerCysAsnLysGlyPheTrpIleLysGlyLysLysSerThrCysGluAlaThr 2189
962 TACAGCTGCACAGAGGGTTTACATCAAGAGGGAAGAGAGAGACCTCCGAGAGCCCTCA 1021
2190 GlyGlnTrpSerSerProIleProThrCysHisProValSerCysGlyGluProProLys 2209
1022 GGGCAGTGGAGTAGTCTTATACGACGTGCCACCCCGTATCTTGTGTGTGAACCACTAAG 1081
2210 ValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValAlaTrp 2229
1082 GTTGAGATGGCTTTCTGAGGACATACACTGGCAGGATCTTTTGAGAGTGAAGTGAAGTAT 1141
2230 GlnCysAsnProGlyTrpLysSerValGlySerProValPheValCysGlnAlaAsnArg 2249
1142 CAGTGTAAACCCGGCTATAGCCAGTCGGAAGTCTCTGTTTGTCTGCCCATGCAATCGC 1201
2250 HisTrpHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysProProPro 2269
1202 CACTGGCACAGTGAATCCCTCTGATGTGGGTTCTCTGACTGTGGANAACCTCCCCCG 1261
2270 IleGluAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePhe 2289
1262 ATCCAGAAATGGCTTATGAAGAGGAGAAAATTTTGAAGTAGGGTCCCAAGGTTTCAAGTTTC 1321
2290 CysAsnGluGlyTrpGluLeuValGlyAspSerSerTrpThrCysGlnLysSerGlyLys 2309
1322 TGTAAATGAGGGTATGAGCTTGTGTGTGACAGTCTTGTGACATGTTCAGAAATCTGGCAAA 1381
2310 TrpAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeu 2329
|||||
```

Db 1382 TGGATAAAGAGTCAATCCAAAGTGCATGCTGTCACCAAGTGCACAGAGCGCCCTCTTGG 1441
Qy
Qy 2330 GluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValThrPheSerCys 2349
Db 1442 GAAACACAGCTAGTATTAAGAGAGTTGACACACGAGTAGGAGTTGTCACATTTCTCTGT 1501
Qy 2350 LysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrp 2369
Db 1502 AAGAAGGAGCATGCTCTGCAAGGCCCTCTGCTCTGAAATGCTTGCAATCCAGCATGG 1561
Qy 2370 AsnAspSerPheProValCysLysIleValLeuCysThrProProLeuIleSerPhe 2389
Db 1562 AATGACTCTTTCTCTGTTGTAAGATTGTTCTTTGTAGCCACCTCCCTCTAAATTTCTCTT 1621
Qy 2390 GlyValProIleProSerSerAlaLeuHisPheCysSerThrValLysThrCysVal 2409
Db 1622 GGTGTGCCCATCTCTCTCTGCTCTCTCAATTTGGAGTACTGTCAAGTATTCTTGTGTA 1681
Qy 2410 GlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSer 2429
Db 1682 GGTGGGTTTTCTTAAGAGGAATTTTACCACCTCTGCAACCTGATGAGCACTGGAGC 1741
Qy 2430 SerProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsnGly 2449
Db 1742 TCTCCACTGCGCAAGATGTTCTCCAGTAGAATGTCCCAACCTGAGGAAATCCCAATGGA 1801
Qy 2450 IleIleAspValGlnGlyLeuAlaThrLeuSerThrAlaLeuThrCysLysProGly 2469
Db 1802 ATCATTTGATGGAGGCGCTTGGCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGC 1861
Qy 2470 PheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGly 2489
Db 1862 TTTGAAATGGTGGGAATATCTACACCTTTGTGGAGAAATGCTCACTGCTTGGAGGA 1921
Qy 2490 LysProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPhe 2509
Db 1922 AAACCAACATGTAAGCCATTGAGTGCTGAAACCCCAAGGAGATTTTGAATGGCAATTC 1981
Qy 2510 SerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArg 2529
Db 1982 TCTTTACAGGACCTACACTATGGACAGACCGTTTACCTACTCTTGCAACCGAGGCTTTCGG 2041
Qy 2530 LeuGluGlyProSerAlaLeuThrCysLeuLysThrGlyAspTrpAspValAspAlaPro 2549
Db 2042 CTGGAAGGTCCTGAGTGCCTTGACCTGTTTAGACACAGGTGATGGAGTAGATGCCCA 2101
Qy 2550 SerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGly 2569
Db 2102 TCTTGCATGCCATCCACTGTGATTCCCAACACCCATTCGAAATGGTTTGTAGAGGT 2161
Qy 2570 AlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAla 2589
Db 2162 GCAGATTACAGCTATGGTGCATAATCATCTACAGTTGCTTCCCTGGGTTTCAGTGGCT 2221
Qy 2590 GlyHisAlaMetGlnThrCysGluGlySerGlyTrpSerSerSerIleProThrCysMet 2609
Db 2222 GGTGATGCCCATGACAGACCTGTGAAGAGTCAAGATGGTCAAGTTCCATCCCAACATGATG 2281
Qy 2610 ProIleAspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAsp 2629
Db 2282 CCAATAGACTGTGGCCCTCCCTCTCATATAGATTTTGTGAGACTGTACTAAACTCAAGAT 2341
Qy 2630 AspGlnGlyTyrPheGluGlnGlnAspAspMetMetGluValProTyrValThrProHis 2649
Db 2342 GACCAAGGATATTTGAGCAAGAGACGACATGATGAAGTTCCATATGTGACTCTCTCAC 2401
Qy 2650 ProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAla 2669
Db 2402 CTCTCTTATCATTTGGAGCAGTGGCTTAACCTGGGAAATACAAAGGAGTCTCTGCT 2461
Qy 2670 ThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyr 2689
Db 2462 ACACATTTCATCAACATTTCTGTATGGTATGATGTTTCAACCTTGTATATCCAGGATAT 2521

Qy 2690 GluLeuLeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAla 2709
Db 2522 GAATCTCTGGGGAACCCCTGTGCTGATCTGCCAGAGATGGAACCTAGGAATGGCAGTGCA 2581
Qy 2710 ProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArg 2729
Db 2582 CCATCTGCAATTTCAATTTGAATGTGACTTGCCTACTGCTCTCTGAAATGGCTTTTGGGT 2641
Qy 2730 PheThrGluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyHisIleLeu 2749
Db 2642 TTTACAGAGACTAGCATGGGAAGTGTGTGTCAGTATAGCTGTAAACCTGACACATTCTA 2701
Qy 2750 AlaGlySerAspLeuArgLeuLeuGluAsnArgLysTrpSerGlyAlaSerProArg 2769
Db 2702 GTGGGCTCTGGCTTAAAGGCTTTGTCTAGAGAAATAGAAGTGAAGTGGTGGCTTCCCAAGC 2761
Qy 2770 CysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySer 2789
Db 2762 TGTGAAGCCATTTTCATGCAAAAGCCAAATCCAGTCAATGAATGCATCCATCAAGGAAGC 2821
Qy 2790 AsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGly 2809
Db 2822 AACTACACATACCTGACACGTTGTACTATGAGTGTGACCCCGGATATGTCTGAATGGC 2881
Qy 2810 ThrGluArgThrCysGlnAspAspLysAsnTrpAspGluAsnGluProIleCysIle 2829
Db 2882 ACTGAGAGGAGAACATCCAGGATGACAAAACTGGGATGAGGATGAGCCCATTTGCAAT 2941
Qy 2830 ProValAspCysSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTyr 2849
Db 2942 CCTGTGACCGCAGTTTCAACCCAGTCTCAGCCAAATGGCCAGGTGAGAGGACGAGTAC 3001
Qy 2850 ThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArg 2869
Db 3002 ACATTCAAAAGAGATTGAATACATTCGAATGAAGGTTCTTGTCTTGAGGAGGCCAGG 3061
Qy 2870 SerArgValCysLeuAlaAsnGlySerTrpSerGlyValThrProAspCysValProVal 2889
Db 3062 AGTGGGTTTGTCTGTGCAATGGAAATGGAGTTGGAGTGGAGCCACTCCCGACTGTGTGCTGT 3121
Qy 2890 ArgCysAlaThrProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPhe 2909
Db 3122 AGATGTGCCACCGCCCAACCTGGCCAAATGGGGTGACGGAAGCCCTGACTATGGCTTC 3181
Qy 2910 MetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeu 2929
Db 3182 ATGAAGGAAGTAAACATTCCTGTGACAGGGCTTACATCTTGCACGGTGTCTCAAACTC 3241
Qy 2930 ThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCys 2949
Db 3242 ACCTGTCAAGTCAAGTGAAGTGGAGTGGAGAGATTCCTCTCTGTAAACCAAGTCAACTGT 3301
Qy 2950 GlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGly 2969
Db 3302 GGACCTCTGAGAGTCTTGGCCATGGTTTCCCTAATGGTTTCTTTATTATTATGGGGGC 3361
Qy 2970 HisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCys 2989
Db 3362 CATATACAGTATCAGTGTCTTCTGTTTAAAGCTCCATGGAATTTCAATCAAGAGGTGC 3421
Qy 2990 LeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThr 3009
Db 3422 CTCTCCAAATGGCTCTCGAGTGGCAGCTCACCTTCTGCTGCTTCCCTTTTCAGATGTTCCAA 3481
Qy 3010 ProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArg 3029
Db 3482 CCAGTAATGAATATGGAACCTGTAATGGGACAGATTTTGACTGTGGAAAGGACGCCCGG 3541
Qy 3030 IleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAsp 3049
Db 3542 ATTCAGTGTCTCAAGGCTTCAAGCTCTTAGGACTTTTCTGAATCAACCTGTGTGAAGCCGAT 3601

Db 193 TATATCTCTGCTGGTCTTGACACCAATTGAATGCTTGGCCGACGCGCAAGTGGAGTAGAAGT 252
Qy
2013 AspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAlaAspHisAlaSerPro 2032
Db
253 GACCAGCAGTGGCTGCTCTCTGATGATGAGCCACCAATGTTGGACCAACGCTCTCCA 312
Qy
2033 GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrCysSerAspGlyTyrSer 2052
Db
313 GAGACTGCCCATCGGCTCTTTGGAGACATTCGANTCTACTACTGCTCTCATGATGTTACAGC 372
Qy
2053 LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGly 2072
Db
373 CTAGCAGACAAATCCAGCTTCTCTGCATGTCGCCAGGCGCAAGTGGGTACCCCCAGAGGT 432
Qy
2073 GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer 2092
Db
433 CAAGACATGCCCGTGTATAGTCTATCTCTGTGAAAAACCTCCATCGGTTTCTATAGC 492
Qy
2093 IleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys 2112
Db
493 ATCTTTGGAACTCTGTGAGCAAGCAAAATTTGCAGCTGGCTCAGTTGTGAGCTTTAAATGC 552
Qy
2113 MetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrp 2132
Db
553 ATGGAGGGTCTTACTGAAACCTCAGCAAGATTTGAATGATGAGAGGTGGGCAAGTGG 612
Qy
2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle 2152
Db
613 AACCCTTCCCCATGCTCCAGTGCATCCCTGTGCGGTGGAGAGCCACCAAGCATC 672
Qy
2153 MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
Db
673 ATGAATGGCTATGCAAGTGGATCAAACTACAGTTTGGAGCCATGTGGCTTACAGCTGC 732
Qy
2173 AsnLysGlyPheTyrIleLysGlyGluLysLysSerThrCysGluAlaThrGlyGlnTrp 2192
Db
733 AACAGGGGTCTACATCAAGGGGAAAGAGAGACCTGCGAAGCCACAGGGCAGTGG 792
Qy
2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
Db
793 AGTAGTCTTATACCGACGTGCCACCGGTATCTTGTGTGAACCACTAAGGTGTAGAAT 852
Qy
2213 GlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGlnCysAsn 2232
Db
853 GGCCTTCTCGAGCATACAACTGCCAGGATCTTTGAGAGTGAAGTAGGTATCAGTGTAA 912
Qy
2233 ProGlyTyrLysSerValGlySerProValPheValCysGlnAlaAsnArgHisTrpHis 2252
Db
913 CCGGCTATAGTCAGTCCGAGTCTCTGTATTGTCTGCCAGCCATCGCCACTGGCAC 972
Qy
2253 SerGluSerProLeuMetCysValProLeuAspCysGlyLysProProProIleGlnAsn 2272
Db
973 AGTGAATCCCTCTGATGTGTCTCTCTCGACTGTGGAACCTCCCGCGATCCAGAAT 1032
Qy
2273 GlyPheMetLysGlyGluAsnGluValGlySerLysValGlnPhePheCysAsnGlu 2292
Db
1033 GGCCTCATGAAGGAGAAACCTTTGAAGTAGGGTCCAAAGGTTCAGTCTTTCTGTAAATGA 1092
Qy
2293 Gly---TyrGluLeuValGlyAspSerTrpThrCysGlnLysSerGlyLysTrpAsn 2311
Db
1093 GGGTTATNGAGCTTTGTTGGGACAGTCTTGGACATGTCAGAAATCTGGCAATGGAAT 1152
Qy
2312 LysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGluAsn 2331
Db
1153 AAGAGTCAAAATCCAAAGTGCATGCCCTGCCAAGTGGCCAGAGCGGCCCTCTTTGGAAAC 1212
Qy
2332 GlnLeuValLysGlyGluLeuThrThrGluValGlyValValThrPheSerCysLysGlu 2351
Db
1213 CAGCTAGTATTAAGAGGTGACACCGAGGTAGGAGTTGTGACATTTCTCTGTAAAGAA 1272
Qy
2352 GlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAsp 2371
Db
1273 AGGCATGCTCTGCNAGGCCCTCTGTCTCTGAAATGCTTGCCATCCCGCAATGGAATGAC 1332

Qy 2372 SerPheProValCysLysIleValLeuCysThrProProLeuIleSerPheGlyVal 2391
Db 1333 TCTTTCCCTGTTGTGAAGATTGTCTTTGTACCCACCTCCCTAAATTTCTTTGGTGTG 1392
Qy 2392 ProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGlyGly 2411
Db 1393 OCCATTCCTCTCTGCTCTTCATTTTGGAGTACTGTCAAGTATTTCTGTGTAGGTGG 1452
Qy 2412 PhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerPro 2431
Db 1453 TTTTTCCTAAGAGAAATTTCTACACCTCTGCAACCTGATGGACCTGGAGCTCTCCA 1512
Qy 2432 LeuProGluCysValProValGluCysProGlnProGluLeuIleProAsnGlyIle 2451
Db 1513 CTGCCAGATGTGTTCCAGTAGAATGTCCCAACCTGGAGAAATCCCAATGGAAATCAT 1572
Qy 2452 AspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPheGlu 2471
Db 1573 GATGTCCAGAGCTTGCTCTATCTCAGCACAGCTCTCTATATCTTGAAGCCAGGCTTGAA 1632
Qy 2472 LeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyPro 2491
Db 1633 TTGGTGGAAATATCTACACCTTTGTGGAGAAATGGTCACTGGCTTGGAGGAAACCA 1692
Qy 2492 ThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyr 2511
Db 1693 ACATGTAAAGCCCATTCAGTGGCTGAAACCCCAAGAGATTTTGAATGGCAATTTCTTTAC 1752
Qy 2512 ThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGlu 2531
Db 1753 ACGACTTACACTATGACAGACCGTTTACTCTCTTGTCAACCGAGGCTTTCGGCTCGAA 1812
Qy 2532 GlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCys 2551
Db 1813 GGTCCAGTGCCTTGACCTGTTTAGAGACAGGTGATTTGGAGTGTAGATGCCCATCTTGC 1872
Qy 2552 AsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAsp 2571
Db 1873 AATGCCATCCATGTGATTTCCCAACCACTTGAATAATGGTTTTGTAGAGGTGGAGAT 1932
Qy 2572 TyrSerTyrGlyAlaIleIleIleTyrSerCysPheProGlyPheGlnValAlaGlyHis 2591
Db 1933 TACAGCTATGGTGCCATATCATCTACAGTTGCTTCCCTGGGTTTCAGGTGGCTGGTAT 1992
Qy 2592 AlaMetGlnThrCysGluGluSerGlyTrpSerSerSerIleProThrCysMetProIle 2611
Db 1993 GCCATGACAGCTGTGAAGAGTCAGGATGGTCAAGTTCCATCCCAACATGTATGCCAATA 2052
Qy 2612 AspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGln 2631
Db 2053 GACTGTGGCTCCCTCTCATATAGATTTTGGAGACTGTACTAACTCAAGATGACCG 2112
Qy 2632 GlyTyrPheGluGlnGluAspMetMetGluValProTyrValThrProHisProPro 2651
Db 2113 GGATATTTGGCAGCAGACGACGATGATGGAAGTTCATATGTGACTCCTCACCTCCT 2172
Qy 2652 TyrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHis 2671
Db 2173 TATCATTTGGAGCAGTGGCTAAACCTGGGAAAAATACAAAGGAGTCTCTCTGTCTACAT 2232
Qy 2672 SerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeu 2691
Db 2233 TCATCAAACTTTCTGTATGGTATACATGGTTCATACACCTGTAACTCAGGATATGAACCT 2292
Qy 2692 LeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySerAlaProSer 2711
Db 2293 CTGGGNAACCTGTGTGTGATCTGCCAGAGATGGAATTTGGAATGGCAGTGCACCATCC 2352
Qy 2712 CysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuArgPheThr 2731
Db 2353 TGCATTTCAATGAATGTGACTTGTCTCTGCTCTCTGAAATGCTTTTTCGTTTTTACA 2412

QY 2732 GluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyHisIleLeuAlaGly 2751
DB 2413 GAGACTAGCATGGGAAGTCTGTGAGTATAGCTGTAAACCTGGACACATTTAGCAGGC 2472
QY 2752 SerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaSerProArgCysGlu 2771
DB 2473 TCTGACTTAAGGCTTTGTCTAGAGATAGAAAGTGAGTGGTGGCTTCCCAACGCTGTGAA 2532
QY 2772 AlaIleSerCysLysProAsnProValMetAsnGlySerIleLysGlySerAsnTyr 2791
DB 2533 GCCATTTCAAGAAAGCAAAATCCAGTCATGAATGGATCCATCAAGGAAGCAACTAC 2592
QY 2792 ThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThrGlu 2811
DB 2593 ACATACCTGAGCACGTTGTACTATGAGTGTGACCCCGGATATGTGCTGAATGGCACTGAG 2652
QY 2812 ArgArgThrCysGlnAspAspLysAsnTrpAspGluAspGluProIleCysAlaProVal 2831
DB 2653 AGGAGAACATGCCAGATGACAAACTGGGATGAGGATGAGCCCATTTGCATTTCTGTG 2712
QY 2832 AspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThrPhe 2851
DB 2713 GACTGCAGTTCAACCCAGTCTCAGCAATGCCAGTGCCAGTGAGAGGAGAGTACACATTC 2772
QY 2852 GlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaArgSerArg 2871
DB 2773 CAAAAAGAGATTGAATACACTTGCAATGAAGGGTTCTTGCTTGAGGGAGCCAGGAGTCGG 2832
QY 2872 ValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValProValArgCys 2891
DB 2833 GTTGTCTTGGCAATGGAAAGTTGGAGTGAGGACCTCCCGACTGTGTCTGTACAGTGT 2892
QY 2892 AlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMetLys 2911
DB 2893 GCCACCCGCCCAACTGGCCAAATGGGGTGCGGAGGCGCTGACTATGGCTTCATGAAG 2952
QY 2912 GluValThrPheHisGlyGluGlyTyrIleLeuHisGlyAlaProLysLeuThrCys 2931
DB 2953 GAAAGTAACATTCACCTGTCCAGGGGCTACATCTTGCAAGGGTCTCCAAAACTCACCTGT 3012
QY 2932 GlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyPro 2951
DB 3013 CAGTCAGATGGCACTGGATGCAGAGATTCCTCTCTGTAAACCACTCACTGTGGACCT 3072
QY 2952 ProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyHisIle 2971
DB 3073 CCTGAAGATCTTGCCCATGTTTCCATAGTGTCTTTCTTTATTCATGGGGGCCATATA 3132
QY 2972 GlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSer 2991
DB 3133 CAGTATCAGTGTCTTCTGTTTATAGCTCCATGGAAATTCATCAAGAGGTGCTCTCC 3192
QY 2992 AsnGlySerTrpSerGlySerProSerCysLeuProCysArgCysSerThrProVal 3011
DB 3193 AATGGCTCTGGAGTGGCAGCTCACCTTCTCTGCTGCTTGGATGTTCACACCACTA 3252
QY 3012 IleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGln 3031
DB 3253 ATTGAATATGAACGTCTCAATGGGACAGATTTTGTCTGTGGAAGGACGCCCGGATTCAG 3312
QY 3032 CysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln 3051
DB 3313 TGTCTTCAAGGGCTTCAAGCTCTTAGACTTTCTGAATCCTCTGTGAAGCCGATGGCCAG 3372
QY 3052 TrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIlePro 3071
DB 3373 TGGAGCTCTGGGTTCCCACTGTGACACACATCTTGTGGTCTCTTCCATGATACCA 3432
QY 3072 AsnAlaPheIleSerGluThrSerTrpLysGluAsnValIleThrTyrSerCysArg 3091
DB 3433 AATGGCTTCATCAGTGAGACCACTCTTGGAAGGAAATGTATAACTTACAGCTGCAGG 3492
QY 3092 SerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSer 3111

DB 3493 TCTGATATGTATCAACAGCGAGTTTCAGATCTGATTTGTACAGAGAAAGGGGTATGAGC 3552
QY 3112 GlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAlaAsnAla 3131
DB 3553 CAGCCCTTATCCAGTCTGTGAGCCCTGTCTCTGTGGTCCCAACCGCTCTGTCCGCAATGCA 3612
QY 3132 ValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGly 3151
DB 3613 GTGGCAACTGGAGAGGCACACACTATGAAAGTGAAGTGAATCTCAGATGTCTGGAAAGT 3672
QY 3152 TyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyValTrpPhePro 3171
DB 3673 TATACGATGATACAGATACAGATACATTCACCTGTGCAGAAAGATGGTCCGTGGTCCCT 3732
QY 3172 GluArgIleSerCysSerProLysCysProLeuProGluAsnIleThrHisIleLeu 3191
DB 3733 GAGAGAACTCTCGAGCTCTAAANAATGTCTCTCTCCCGGAAAAACATAACACATATCTT 3792
QY 3192 ValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyr 3211
DB 3793 GTACATGGGAGCAATTCAGTGTGAATAGGCAAGTTTCTGTCTCATGTGCAAGAGGGTAT 3852
QY 3212 ThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPhe 3231
DB 3853 ACCTTTGAGGGAGTTAAACATATCAGTATGTACAGCTTGATGGAACTTGGAGGCCCATTC 3312
QY 3232 SerAspGluSerCysSerProValSerCysGlyLysProGluSerProGluHisGlyPhe 3251
DB 3913 TCCGATGATCTTCAGTCCAGTTCTTGTGGGAAACCTGAAAGTCCAGAACATGATTT 3272
QY 3252 ValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyr 3271
DB 3973 GTGGTTGGCAGTAAATACACCTTTGAAAGCACAAATTAATTTATCAGTGTGAGCCTGCTAT 4032
QY 3272 GluLeuGlu----- 3274
DB 4033 GAACTAGAGAAATTTGGCTGTGTAATCCATCTGTCTCTGGACTTTTCTTGGTTGACAGGACC 4092
QY 3274 ----- 3274
DB 4093 CTCAGCTGCAGTGGAGTTGGCTAGAGTCCAAATCCAGACCTGTGTGCTGGGTATCA 4152
QY 3274 ----- 3274
DB 4153 GCAGCAGAGGGTGCAGAAACAGCGGATATTTGGTGAACCGCAAAATGCTGCTGCTGATCAT 4212
QY 3274 ----- 3274
DB 4213 CCTCTGGAAGTTTGTCTCAGAGAGATATCCCGGCCATGTGAGGTGTCACTCCGCCCTAC 4272
QY 3275 -----GlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAla 4292
DB 4273 TGGGGGGGGGAAACAGGGAACTGTCTGCCAGGAGAAACAGACAGTGGAGTGGAGGGTGGCA 4332
QY 3293 IleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLysAlaAspIle 3312
DB 4333 ATATGCAAGAGACAGAGGTGTAATCTCCACTTGAATTTCTCAATGGGAAGCTGACAT 4392
QY 3313 GluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGlu 3332
DB 4393 GAAACAGAGACGACTGGACCCCAACCTGGTATATTTCTTGCACACAGAGGCTACAGTCTTGA 4452
QY 3333 GlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCys 3352
DB 4453 GGGCCATCTGAGGCACACTGCACAGAAAATGAACTCGAGGCCACCCAGTCCCTCTCTGC 4512
QY 3353 LysProAsnProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLys 3372
DB 4513 AAACCAATCCATGCTGCTTCTTTGTGATTCGAGAAATGCTCTGCTCTGAAAG 4572
QY 3373 GluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGly 3392

Db 4573 GAGTTTATGTTGATCAGATGTGTCCATCAATGTAGGGAAGGTTTCTGCTGCAGGGC 4632
Qy 3393 HisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGlu 3412
Db 4633 CAGGGCATCAATACCTGCAACCCCGACGACGACGCTGGACACAGACGCGCAAAATGTGA 4692
Qy 3413 Lys----- 3413
Db 4693 AGAAGATATACACACAGCCCAAGTCCCTCGAATTTTCAGCTAGCAGCTTATTCAGTATT 4752
Qy 3413 ----- 3413
Db 4753 AGAATGTTTATTTTGGGGGAGGGGTTCAAGATGGCCAACTAGAAACAGCTGTGGCCGGA 4812
Qy 3413 ----- 3413
Db 4813 GCCTCCACCGAGAGAACAAACAAACAAAGCGAGAAAGCGTGTGTACAAAGGCCCA 4872
Qy 3413 ----- 3413
Db 4873 CTTGGGAGGCACATGTGGGGCAAGCGAGCTCCACCCCGCCAAAGAGGAGGTGGACCTCC 4932
Qy 3413 ----- 3413
Db 4933 TGGGGGAAATTCAGCAACTCCAGCCAGGGGTTTATGAACAGACCTCTGATCTCCCTCGA 4992
Qy 3413 ----- 3413
Db 5053 CTTGCTGGCTCTGAGGAATCCAGCGAGCTGGACTAGTGGGATTCCTCCACAGCACAGTTT 5112
Qy 3413 ----- 3413
Db 5113 ACCTGCTCTGCCAAGGGCAGCTAGAGCGCTTGTGTAAGCGAGTCCCTGATCCCATGCC 5172
Qy 3413 ----- 3413
Db 5173 CTTGATTGGGATGAGACCCCCCAACACAGGGGTCAAGATGAGACCCCCCAACACAGG 5232
Qy 3413 ----- 3413
Db 5233 GGTACACAGACACTTATACAGGGTGTCTCTGTAGCATCAGTCACTAGTGCCTCTGGG 5292
Qy 3413 ----- 3413
Db 5293 ACAGAGCTCCACAGAGAAAGCAGCAGCGCCATCTTTGCTGTTCTGCGAGGTCCGCTGGA 5352
Qy 3413 ----- 3413
Db 5353 AAAGCACAGAAATGGCGCAGAGGCTAGGATGTGATGTAATTTGAAGAGTAGGCTTCAGAAAG 5412
Qy 3413 ----- 3413
Db 5413 TGGGTAATATGAGTTTCGCTGAGCTAAAGGAACATGTTCTTAACCAATGCAAGAGCC 5472
Qy 3413 ----- 3413
Db 5473 AAGAACCGAGTAATAACATTACAGGATCCGTTAACAGATAACCAAGTTTAGAAAGGAAT 5532
Qy 3413 ----- 3413
Db 5533 GTAATGACCTGATGGAGCTGAAACAAACACAGAGAACTTCAATGCAACAAACAAA 5592
Qy 3413 ----- 3413
Db 5593 CAAGGCCAATCTCCAGTTCCAGGAATCCAGAGAACCCCGAGTAAGTACTTCCATGAGAAG 5652
Qy 3414 ----- 3420
Db 5653 ATCAACCCCAAGACACATAATCTCAGGTTCTCCAGAAATCTCATGTGTGTCACCAAGCT 5712

Qy 3421 HisValGluAsnAlaIleAlaArgGlyValHisIleThrGlnIleThrValAspMetIleThrTyr 3440
Db 5713 CAGTAGAATAATGCAATTTCTCGAGGCGTACATATCATATGAGACATGATCACTTAC 5772
Qy 3441 SerCysTyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGly 3460
Db 5773 TCATGTTACAGTGGATACATGTTGGAGGGTTTCTCGAGGAGTGTGTTGTTAGAAATGGA 5832
Qy 3461 ThrTrpThrSerProProlleCysArgAlaValCysArgPheProCysGlnAsnGlyGly 3480
Db 5833 ACATGGACATCACTCTCTATTTCAGAGCTGTCTGTCTGATTTCTCAGATTCAGATGGGGC 5892
Qy 3481 IleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGlu 3500
Db 5893 ATTCGCCAACGCCCAATGCTTGTCTCTCAGAGGCTGATGGGGCGCTCTCTGGA 5952
Qy 3501 GluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCys 3520
Db 5953 GAACCAATCTGCATCTTCTCTGTCGACGAGGTGCTGTGTGGCCCTTACCAGTGT 6012
Qy 3521 AspCysProProGlyTrpThrGlySerArgCysHisThr 3533
Db 6013 GACTGCCCGCTGCTGCGAGCGGTCTCTGCTGTCATACA 6051

RESULT 11
LOCUS HSM806032 4701 bp mRNA linear PRI 17-JUN-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp686110142 (from clone DKFZp686110142).
ACCESSION BX537918
VERSION BX537918.1 GI:31873946
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4701)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Pobo, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686110142) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES
source 1. 4701
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="9q31.2-32"
/clone="DKFZp686110142"
/tissue_type="human colon endothel primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1-Sti; host DH10B; sites SfiI + SfiIIB"
/dev_stage="adult"
1. 4701
/gene="DKFZp686110142"
1. 4326
/gene="DKFZp686110142"
/notes="similar to polydom protein (Mus musculus), N-terminus truncated"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD97901.1"
/db_xref="GI:31873947"

/translation="QNNPSPMSIQIPVRCGEPPSPIMNGVAGSNVSPFCAMVAVSNK
 GYIEBKKATQNSSPIDTCHPVSCGEPKPVNGFLEHTGRIFESVRQCN
 POKYSGSVFVQNRHWHSSPLMCPVLDCKGPPPIQNGFMKGFVGVSKVQFC
 NBYELVGSSTVCKSGKWNKSNPKPAKCPBPPLLENQLVLKBLTFRVGVTFK
 CKEHVLQGPSVLCPLPSQOMNSDFVCKIIVLCCTPPLISFGVPISSALHFGSTVKY
 SCVGGPFLRGNTTLLCGPDGWTSSPLPECPVECPQPBEPNGIIDVOQLAYLSTALY
 TCKKPELNGNTTLLCGPDGWTSSPLPECPVECPQPBEPNGIIDVOQLAYLSTALY
 TCKKPELNGNTTLLCGPDGWTSSPLPECPVECPQPBEPNGIIDVOQLAYLSTALY
 SCNRGRLSGPRLTCLGPDGWTSSPLPECPVECPQPBEPNGIIDVOQLAYLSTALY
 SCNPGVAVAGHMTCEBSSSSIPICMPIDGCLPPLHIDGCTCKLDDQGYFQED
 DMWEPVYVPPHYPHILGAVAKWENTKESPATHSSNPLGYTWVSYTCNPGYELGNPV
 LIQEDGTWNGSAPSCISLTCOLPAPENGFLRTFTSMGSAVOYSCKPGHLAGSDL
 RLCLERNKWSGAPRCEISLCKPNPVMNGSIKGSNTYLSLTYEBCDPYVLLGTER
 RTCODKWNDEDEPICIPVDCSPVUSANGOVGRDEYTPQKEIHTCNBGFLLGARS
 RVCLANGSAGATPDVPCVRCATPPOLANGVEGLDYGFPMKEVTFCHBGYILHGNPK
 LTCSDGNWDAIPLCKPVCNCPEDLAHGFNGFYIHGGHIQYGFPGYKLLGNSS
 RRLCLNGWSGSPCLCRCSTPVLEYGTGTDPCGKAARIQCFPGFLGSLSEI
 TCBAQMSGSPHCHTSQGLPMIENAFISETSMKENVITVSCSGVYVQGSDDL
 ICTBKGVNSQPVPCBPLSGSPSVANAVATGEATYSEVKRLCLEGVYMDTDT
 ITCKDGRWPPERIISCSPKKCPLENI THILVHGDGFSVNRQVSUSCAGYTFKGVNI
 RERVCOENKONGVAICKECTCEBTPLEFLNGKADIBNRTTGNVNVYSNCRGYSLEGP
 SRAHCTENGWGHVPLCKPFCVPPVI PENALLSEKFPYVDQNVYSICNREGFLQ
 HGIITCNPDETWTSAPCEKISCGPAPNGALARGVHYQYQDMITVSCYGYMLEG
 FLRSVCLNGSTWTSPIICRAVCFPCQNGIICORPNACSCPEGMGLCEBPICILPC
 LKGRCAVAPYQDCPQGTGWSRCHTAVCQSPCLNGSKVRPRCHCLSSWTHGNSRK
 RRTGP"

polya_signal

4663..4668

/gene="DKFZp686110142"

polya_site

4680

/gene="DKFZp686110142"

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	4701
Score:	8157.00	Matches:	1437
Percent Similarity:	99.86%	Conservative:	2
Best Local Similarity:	99.72%	Mismatches:	2
Query Match:	40.84%	Indels:	0
DB:	9	Gaps:	0

US-09-977-053-4 (1-3571) x HSM806032 (1-4701)

QY	2131	GlnTrpAsnProSerProMetSerIleGlnCysIleProValArgCysGluProPro	2150
DB	1	CAGTGAACCCCTTCCCAAGTCCATCCAGTGCATCCCTGTGGGGTGTGGAGGCCACCA	60
QY	2151	SerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyr	2170
DB	61	AGCATCATGATGGCTATCGAAGTGCATCAACTACAGTCTTTGGAGCCATGGTGGCTTAC	120
QY	2171	SerCysAsnLysGlyPheTyrIleLysGlyGluLysSerThrCysGluAlaThrGly	2190
DB	121	AGCTGCAACAAGGGTCTTACATCGAAGGGGAAAGAGAGCACCCTGCGAAGCCACAGGG	180
QY	2191	GlnTrpSerSerProIleProThrCysHisProValSerCysGlyGluProProLysVal	2210
DB	181	CAGTGAGTAGTCTCTATACCGAGTGCACCCCGGTATCTTGTGGTGAACCCATCAAGTT	240
QY	2211	GluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGln	2230
DB	241	GAGATGGCTTCTCGAGCATCAACTGGCAGGATCTTTGAGAGTGAAGTGAAGTATCAG	300
QY	2231	CysAsnProGlyTyrIlySerValGlySerProValPheValCysGlnAlaAsnArgHis	2250
DB	301	TGTAAACCCGGGCTATTAAGTCAGTCGGAAGTCTGTATTTGTCTGCGCAAGCCATCGCCAC	360
QY	2251	TrpHisSerGluSerProLeuMetCysValProLeuAsnCysGlyLysProProProIle	2270
DB	361	TGGCAGAGTGAATCCCTCTGATGTGTCTCTCTGACTGTGGAAACCTTCCCCCGATC	420
QY	2271	GlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhePheCys	2290
DB	421	CAGATGGCTCTCATGAAGAGGAAACTTTGAAGTAGGGTCCAGAGTTTTCATTTTCTGT	480

QY	2291	AsnGluGlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGlyLysTyr	2310
DB	481	AATGAGGGTTATGAGCTTGTGTGTGACAGTCTTGTGGACATGTTCAGAAATCTGGCAATGG	540
QY	2311	AsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeuLeuGlu	2330
DB	541	ATAAGAGTCAANTCAAGTGATGCTGCCAGTGCAGAGCGCCCTCTTGAA	600
QY	2331	AsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSerCysLys	2350
DB	601	AACCACTAGTATTAAAGGAGTGTGACACCGAGGTAGGAGTGTGACATTTCTCTGAAA	660
QY	2351	GluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGlnTyrAsn	2370
DB	661	GAAAGGCACTGCTCGAAGGCCCTCTGTCCTGAAATGCTTGGCATCCAGCAATGAAT	720
QY	2371	AspSerPheProValCysLysIleValLeuCysThrProProLeuLysSerPheGly	2390
DB	721	GACTCTTCCCTGTTGTAAAGATTGTTCTTTGTACCCCACTCCCTTAATTTCTTTGGT	780
QY	2391	ValProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCysValGly	2410
DB	781	GTCCCATTTCTTCT	840
QY	2411	GlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTyrSerSer	2430
DB	841	GGGTTTTTCTTAAGAGGAAATTTCTACCACTCTGCGCACTCATGTGGCCTCTGGAGCTCT	900
QY	2431	ProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIle	2450
DB	901	CCACTGCCAGATGTTGTTCCAGTAGAATGTCCCACTGCCAGGAAATCCCAATGGAAATC	960
QY	2451	IleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysProGlyPhe	2470
DB	961	ATTGATGTGCAAGGCTTGGCTATCTCAGCACAGCTCTCTATACCTGCAAGCCAGGCTTT	1020
QY	2471	GluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisThrLeuGlyGlyLys	2490
DB	1021	GAAATGGTGGGAAATACTACCACTTGTGGAGAAATAGGTCACTGGCTTGGAGGAAAT	1080
QY	2491	ProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLysPheSer	2510
DB	1081	CCACATGTTAAGCCATTTGAGTGCCTGAAACCCAGAGGAGATTTGAATGCAAAATCTCT	1140
QY	2511	TyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeu	2530
DB	1141	TACACGACCTACACTATGACACAGATCGTTACCTCTCTTGCACCCAGAGGCTTTGGCTC	1200
QY	2531	GluGlyProSerAlaLeuThrCysLeuGluThrGlyAspThrAspValAlaProSer	2550
DB	1201	GAAAGTCCAGTGCCTTGGACCTGTTTAGAGACAGGTGATTTGGATGTAGATGCCCATCT	1260
QY	2551	CysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGluGlyAla	2570
DB	1261	TGCATGCTCATCTGCTGATTTCCCAACCAACCATTTGMAAATGTTTGTAGAAGGTGCA	1320
QY	2571	AspTyrSerTyrGlyAlaIleIleLysSerCysPheProGlyPheGlnValAlaGly	2590
DB	1321	GATTACAGCTATGGTGCCTAATCATCATCAGTGTCTTCCCTGGGTTTTCAGGTGGCTGGT	1380
QY	2591	HisAlaMetGlnThrCysGluSerGlyTyrProSerSerIleProThrCysMetPro	2610
DB	1381	CATGCCATGACAGCTGTGAGAGTGTGAGATGTGCAAGTGTCAAGTGTTCATCCACATGTATGCCA	1440
QY	2611	IleAspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLysAspAsp	2630
DB	1441	ATAGACTGTGGCTCTCTCTCTCATATAGATTTTGGAGACTGTACTAACTCAAAGATGAC	1500
QY	2631	GlnGlyTyrPheGluGlnGluAspMetMetGluValProTyrValThrProHisPro	2650
DB	1501	CAGGATATTTTGGAGCAGAGACGATGATGAGAGTTTCCATATGTGACTCTCTCACCTT	1560

Qy	2651	ProTyrHisLeuGlyAlaValAlaLysThrTTrpGluAsnThrIysGluSerProAlaThr	2670
Db	1561	CCCTTATCATTTGGGAGCAGTGGCTTAAACCTGGGAAATAACAAGAGAGTCTCTCGTGTACA	1620
Qy	2671	HisSerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGlu	2690
Db	1621	CATTTCATCAACACTTTCGTATGGTACCTGGTTTCATACACCTGTATTCGGATATGAA	1680
Qy	2691	LeuLeuGlyAsnProValLeuLeuLeuCysGlnGluAspGlyThrTrpAsnGlySerAlaPro	2710
Db	1681	CTTCTGGGGAACCTGTGTGCTGATCTGCGCAGAGATGAACTTGGAAATGCGAGTGCACCA	1740
Qy	2711	SerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeuAtrgPhe	2730
Db	1741	TCCTGCATTTCAATTGAATGTGACTTTCCTGCTACTGCTCTCGAAATGCGCTTTTTCGCGTTT	1800
Qy	2731	ThrGluThrSerMetGlySerAlaValGlnTyrSerCysIysProGlyHisIleLeuAla	2750
Db	1801	ACAGAGACTAGCATGGGAAGTCTGTGCAGTATAGCTGTAAACTGGACACATTCAGCA	1860
Qy	2751	GlySerAspLeuArgLeuCysLeuGluAsnArgIysTrpSerGlyAlaSerProArgCys	2770
Db	1861	GGCTCTGCATTTAGGCTTTGTCTAGAGATAGAAAGTGGAGTGGTGCTCTCCCAAGCTGT	1920
Qy	2771	GluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGlySerAsn	2790
Db	1921	GAAGCCATTTTCATCAAAAAAGCCAAATCCAGTCAATGAATGCATCATCAAGGAAGCAAC	1980
Qy	2791	TyrThrTyrIleuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsnGlyThr	2810
Db	1981	TACACATACCTAGACAAGTTGTACTATAGTGTGACCCCGGATATGTCTGGAATGGGCACT	2040
Qy	2811	GluArgArgThrCysGlnAspAspLysAsnTrpAspGluAspGluProIleCysIlePro	2830
Db	2041	GAGAGGAGACATCCAGGATGACAAATACTGGGATGAGATGAGGCCCATTTGCATTCCT	2100
Qy	2831	ValAspCysSerSerProProValSerAlaAsnGlyGlnValArgGlyAspGluTyrThr	2850
Db	2101	GTGGACTGCGAGTTTCACTCCCTCCAGTCTCAGCCAAATGGCCAGGTGAGAGGAGACGAGTACACA	2160
Qy	2851	PheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAlaAtrgSer	2870
Db	2161	TTCCAAAAAGAGATGGAACACACTTGTCAATGAAGGGTTCTTGTCTTGAAGGAGCCAGGAGT	2220
Qy	2871	ArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValProValArg	2890
Db	2221	CGGGTTGTCTTGGCAATGGAGTTGGATGGAGGCCACTCTCCGACTGTGTGCTGTGACGA	2280
Qy	2891	CysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspTyrGlyPheMet	2910
Db	2281	TGTGCCACCCCGCCACAACTGGCCAAATGGGGGTGACGGAAGCCCTGGACTATGGCTTCATG	2340
Qy	2911	LysGluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProLysLeuThr	2930
Db	2341	AAGGAAGTAACTTCCACTGTTCATGAGGGCTTACATCTTGACCGGTGTCTCCAAACTCACC	2400
Qy	2931	CysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGly	2950
Db	2401	TGTCACTCAGATGCGCAACTGGATGTCAGAGATTCTCTCTGTAAACCAAGTCNACTGTGGA	2460
Qy	2951	ProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHis	2970
Db	2461	CCTCCTGAAGATCTTGCCCATGGTTTCTCCCTAAATGGTTTCTTCTTATTCATGGGGGCGAT	2520
Qy	2971	IleGlnTyrGlnCysPheProGlyTyrIleValLeuHisGlyAsnSerSerArgAtrgCysLeu	2990
Db	2521	ATACAGATATCAGTCTCTTCCGGTTATATAGCTTCCATGGAAATTCATCAAGAGAGTGCCTC	2580
Qy	2991	SerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSerThrPro	3010
Db	2581	TCCAAATGGCTCTCGAGTGGCGAGCTCACCTTCTCTGCTGCTTGCAGATGTTTCCACACCA	2640
Qy	3011	ValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyTysAlaAlaArgIle	3030

Db	2641	GTAAATGAAATATGGAATCTGCAATGGGACAGATTTTGACTGTGGAAGGACGCCGGANT	2700
Qy	3031	GlnCysPheIysGlyPheIysLeuLeuGlyLeuSerGluIleThrCysGluIuIaAspGly	3050
Db	2701	CAGTGCCTCAAGGCTTCAGACTCTAGGACTTTCTGAAATCACCCTGTGAAGCCGATGGC	2760
Qy	3051	GlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIle	3070
Db	2761	CAGTGGAGCTCTGGGTTCCCCACCTGTGAACACACATCTTGTGTGTTCTCTTCCAATGATA	2820
Qy	3071	ProMetAlaPheIleSerGluThrSerSerTrpIysGluAenValIleThrTrpSerCys	3090
Db	2821	CCAAATCGTTCATCAGTGAGACGAGCTCTTGAAGGAAATGTGATACTTACAGCTGC	2880
Qy	3091	ArgSerGlyTrpValIleGlnGlySerSerAspLeuIleCysThrGluIysGlyValTrp	3110
Db	2881	AGGTCTGGATATGTCATACAGGCAGTTTCAGATCTGATTTGTACAGAGAAAGGGTATGG	2940
Qy	3111	SerGlnProTrpProValCysGluProLeuSerCysGlySerProProSerValAlaAsn	3130
Db	2941	AGCCAGCGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCCCACCGTCTGTGCCCAAT	3000
Qy	3131	AlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValIysLeuAryCysLeuGlu	3150
Db	3001	GCAGTGGCACTTGAGAGGCACACACTATGAAAGTGAAGTGAACATCAGATGTCGTGAA	3060
Qy	3151	GlyTrpThrMetAspThrAspThrAspThrPheThrCysGlnIysAspGlyAryTrpPhe	3170
Db	3061	GGTATACGATGATACAGATACAGATACAAATCACCTGTGAGAAAGATGTCGTGCTTC	3120
Qy	3171	ProGluArgIleSerCysSerProIyGlyCysProLeuProGluAenIleThrHisIle	3190
Db	3121	CCTGAGAGAAATCTCCTCAGTCTCTAAAAAATGTCTCTCCCGGAAAAACATAACACATATA	3180
Qy	3191	LeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGly	3210
Db	3181	CTTTGATACATGGGGACGATTTTCAGTGTGAATAGGCAAGTTTCTGTGTCTATGTGCGAAGAGG	3240
Qy	3211	TyrThrPheGluGlyValAenIleSerValCysGlnLeuAspGlyThrTrpGluProPro	3230
Db	3241	TATACCTTTGAGGGAGTTACATATACATATGATGTGATGTGATGGAACCTGGAGCCACCA	3300
Qy	3231	PheSerAspGluSerCysSerProValSerCysGlyIysProGluSerProGluHisGly	3250
Db	3301	TTCTCCGATGAATCTTCAGTCCAGTGTCTGTGGGAAACCTGAAAGTCCAGAACATGGA	3360
Qy	3251	PheValValGlySerIysTyrThrPheGluSerThrIleIleTyrGlnCysGluProGly	3270
Db	3361	TTTGTGTGTGGCAGTAAATACACCTTTGAAAGCAACAATATTTATCATGTGTGAGCCCTGGC	3420
Qy	3271	TyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAenArgGlnTrpSerGlyGly	3290
Db	3421	TATGAACTAGGGGGACAGGGAAGTGTCTGCCAGAGGAACACAGACAGTGGAGTGGAGGG	3480
Qy	3291	ValAlaIleCysIysGluThrArgCysGluThrProLeuGluPheLeuAenGlyIysAla	3310
Db	3481	GTGGCAATATGCAAAAGAGACAGGAGTGAACCTCCATTTGAAATTTCTCAATGGCAAGCT	3540
Qy	3311	AspIleGluAenArgThrGlyProAsnValValTyrSerCysAenArgGlyTrpSer	3330
Db	3541	GACATTTGAAACACGACGATGGACCCCAACGTGGTATATTCCTGCAACAGAGGCTACAGT	3600
Qy	3331	LeuGluGlyProSerGluAlaHisCysThrGluAenGlyThrTrpSerHisProValPro	3350
Db	3601	CTTGAGGGCCNTCTGAGGCACACTGCACAGAAATGGAACCTGGAGCCACCCAGTCCCT	3660
Qy	3351	LeuCysIysProAsnProCysProValProPheValIleProGluAenAlaLeuLeuSer	3370
Db	3661	CTCTGCAAAACCAATCCATGSCCTTGTCTCTTTGTGATTCGAGAGATGCTCTCTGTCTCT	3720
Qy	3371	GluIysGluPheTrpValAspGlnAenValSerIleIysCysArgGluGlyPheLeuLeu	3390

```
Db 3721 GAAAAGGAGTTTATGTTGATCAGAAATGTGTCCATCAAAATGTAGGAAGTTTCTGCTG 3780
Qy 3391 GlnGlyHisGlyLeuThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLys 3410
Db 3781 CAGGGCCACGGCATCATATCACTGCAACCCGACGAGACGTGGACACAGCAAGCGCCAA 3840
Qy 3411 CysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyVal 3430
Db 3841 TGTGAAAANAATCTCATGTGTCCACAGCTCAGGTAGAAAATGCAATGCTCTGAGCGTA 3900
Qy 3431 HisTrpGlnTrpGlyAspMetIleThrTrpSerCysTrpSerGlyTrpMetLeuGluGly 3450
Db 3901 CATTAATCAATGAGACATGATCACTACTCATGTTACAGTGGATACATGTTGAGGGT 3960
Qy 3451 PheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAla 3470
Db 3961 TTCCTGAGGAGTGTGTTGTTAGAAAATGGAACATGACATCACTCTTATTTGACAGGCT 4020
Qy 3471 ValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCys 3490
Db 4021 GTCTGTGATTTCCATGTCAGATGGGGCATCTGCCAACGCCCAATGCTTGTCTCTGT 4080
Qy 3491 ProGluGlyTrpMetGlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsn 3510
Db 4081 CCAGAGGCTGAGTGGGGCGCTCTGTGAAGAACCAATCTGCAATCTTCTCTGTAAC 4140
Qy 3511 GlyGlyArgCysValAlaProTrpGlnCysAspCysProProGlyTrpThrGlySerArg 3530
Db 4141 GGAGGTGCTGTGTGGCCCTTACCAGTGTGACTGCCGCCCTGAGTGGAGCGGGTCTGCG 4200
Qy 3531 CysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsn 3550
Db 4201 TGTATACAGCTGTGTTGCCAGTCTCCCTGCTTAAATGGTGGAAATGTGTAAAGACCAAC 4260
Qy 3551 ArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGly 3570
Db 4261 CGATGTCAGTGTCTTCTCTTGGACGGGACATACACTGTTCAGGAAAGGAGGACTGG 4320
Qy 3571 Phe 3571
Db 4321 TTT 4323

RESULT 12
AX880905 AX880905 5124 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 15810 from Patent EP1074617.
DEFINITION AX880905
ACCESSION AX880905
VERSION AX880905.1 GI:40035641
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patient: BP 1074617-A 15810 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..5124
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
119..4069
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE91130.1"
/db_xref="GI:40035642"
/translation="MASTPKBEHCYLLHSFEEFALRALRHEDLPSSGSPFQDDMVHC
SYLDCGDKCCDRMGSGTGTGHFECICRKYGKGLQYEATCPGTYKPEGSPG
GISSCIPCDNHTSPGSGTSPDCVCEGRASQGTCELVHCAPKPENGYFQNT
CNNHFNACGVRCRHPGDLVSSIIILCLPNGLWSGLSEYCRVRTCPLRQPKGHISC
```

```
STREMLYKTTCLVACDEGYRLGSDKLTCQGNQSDMDGPBPRCVERHCSTFQMPKQVLI
SPHNGXQPAKEGTILCYSCROGFIILSGVKEMLRCTTSGKNVGVCAAVKQVDEAPQI
NCPRDIEARTLEQQDSANTWQIPTAKDNGSKYSHVHPAFTPTPLPGDVALIYVT
ATDLSNGASCIPIHKVDAEPVVDWCRSPPPQVSEKVVHAAWDEPQFSDNSGABL
VITRSHGTQDLFPQGETIIVQYATDPSGNRTCDIHIIVKGSFCBIPTFPVNGDFICT
PMTGVNCTLTLCLEGYDFTEGSDTKYICAYEDGVWKPVTYTTTWDPCBENLTKKYCLEYN
PFMYKAARCDTDLAKQFSAEPTTLGRWVSPSCDABDIWRLEBENLTKKYCLEYN
LYENGFALPGCKWGNRLDYSYDDPLTTOETATISGNKSSRIKRSAPLSYKIK
LLENFASVPLDERDNLWENQBLQTLTITNKLKTLNKDKMPSYQPLASELI
ADNSLGTAKKAPFCRPGSVLRMCNCPGLTYINLEHFTCSRCISQDSBEGQUR
CKLCPGMYETIHSRNIISDCRAQCKQTYSCSGLTSCPLGTTPKPFGRSLC
PENTSTVRKAVNIISACVPCPEKFSRGLMPCHPCPRDYQPNAGKACPLACPFCHS
TTPPAGRSITRCSSTFAAESVVPASLGHIKRHEISSQVFEHCFCFNPCHNS
GTQOOLGRGYVCLPLGYTKACETDIDRCSPLCLNNGVCKDLQVEHCFCPSGYTG
QACENINECSSSPCLNKGIQVGVAGYRCTCVKGFVGLHCETEVNECOSNANAV
CBQVGGFLCKCPGFLGTRCKNVDCLSQKNGATCKDGNANSFRCLCAAGPTGSH
CELNINQCSNPRQATCVDBLNSYCKQPFSGKRCETQSTQFNLDIVSGYIG
YVMLVGLPISLHALCTTWMKSSDDMYGTPTISYAVDNGSDNLTLLTDYNGWLYNG
REKINTCPSVNDGRWHIAITWTSANGIMKVYIDGKLSDGAGLSVGLPITGMP"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 5124
Score: 7537.00 Matches: 1347
Percent Similarity: 99.63% Conservative: 1
Best Local Similarity: 99.54% Minmatches: 5
Query Match: 37.74% Indels: 0
DB: 6 Gaps: 0

US-09-977-053-4 (1-3571) x AX880905 (1-5124)
Qy 195 ThrAspGlyTrpSerAsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSer 214
Db 2 ACTGATGATATATTCAAATGGGGGAGACCTTAGACCAATTCGACGCTACCTGCGAGATTCA 61
Qy 215 GlyValGluIlePheThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMet 234
Db 62 GGAGTGGAGATCTTCACTTTTGGCATATGCGAAGGGAACATTCGACAGCTGTAATGACATG 121
Qy 235 AlaSerThrProLysGluGluHisCysTrpLeuLeuHisSerPheGluGluPheGluAla 254
Db 122 GCTTCCACCCCAAGGAGGAGCACTGTTACCTGCTACACAGTTTGAAGAATTTGAGGCT 181
Qy 255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMet 274
Db 182 TTAGTCTCGCCGGGCAATGCAAGATCTACCTTCTGGGAGTTTATTCAGATGATATG 241
Qy 275 ValHisCysSerTrpLeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCys 294
Db 242 GTCCACTGCTCATATCTTTGTGATGAGGCAAGGACTGCTGTGACCGAATGCGAAGCTGC 301
Qy 295 LysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTrpTrpGlyLys 314
Db 302 AAATGTGGGACACACACAGGCCATTTTGTAGTGTGATCTGTGAAAAGGGGTATTATACGGGAA 361
Qy 315 GlyLeuGlnTrpGluCysThrAlaCysProSerGlyThrTrpLysProGluGlySerPro 334
Db 362 GGTCTGCAATGAAATGACAGCTTGCCCATCGGGACATACAAACTGAGGCTCACCA 421
Qy 335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGlySer 354
Db 422 GGAGGAATCAGCAGTTGCAATTCATGCTGCTGATGAAAATCACACCTCTCCACCTGGAAGC 481
Qy 355 ThrSerProGluAspCysValCysArgGluGlyTrpArgAlaSerGlyGlnThrCysGlu 374
Db 482 ACATCCCTCGAAGACTGTGTCTGACAGAGGAGATACAGGGGCATCTCGGCCAGAGCTGTGAA 541
Qy 375 LeuValHisCysProAlaLeuLysProProGluAsnGlyTrpPheIleGlnAsnThrCys 394
Db 542 CTGTGCTCACTGCCCTGCCCTGAAGCTCCGAAATGTTACTTTATCCAAAACACTTGC 601
Qy 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
```


182 TTAGCTCGCGGGCATTTGCATGAAGATCTACCTCTCGGGAGTTTATTTCAAGATGATATG 241
183 |||||
275 ValHisCysSerThrLeuValCysAspGluGlyValAspCysCysAspArgMetGlySerCys 294
276 |||||
242 GTCCACTGCTCATATCTTTGTGTATGAAGGCAAGAGCTGCTGTGACCAATGGGAAGCTGC 301
243 |||||
295 LysCysGlyThrHisThrGlyHisPheGluCysGluCysGlyValGlyValGlyVal 314
296 |||||
302 AAATGTGGGACACACACAGGCCATTTTGAAGTGCATCTGTGAAAGGGTATTAACGGGAA 361
303 |||||
315 GlyLeuGlnThrGluCysThrAlaCysProSerGlyThrThrLysProGluGlySerPro 334
316 |||||
362 GGTCTGCAGTATGAATGACAGCTTGGCCATCGGGGACATACAAACCTGAAGGCTCACCA 421
363 |||||
335 GlyGlyLeuSerCysValProCysProAspGluAsnHisThrSerProGlyValSer 354
336 |||||
422 GAGGAAATCAGCGTTGCAATCTCATGTCTCTGATGAANAATCAACCTCTCCACCTGGAGC 481
423 |||||
355 ThrSerProGluAspCysValCysArgGluGlyValArgAlaSerGlyGlnThrCysGlu 374
356 |||||
482 ACATCCCTCGAAGACTGTGTCTGCAGAGAGGATACAGGGCATCTGGCCAGAGCTGTGAA 541
483 |||||
375 LeuValHisCysProAlaLeuValCysProGluAsnGlyThrPheLeuGlnAsnThrCys 394
376 |||||
542 CTTGTCCACTGCTGCTGCTGAGGCTCCCGAAATGGTTACTTTATTCAAAACACCTTGC 601
543 |||||
395 AsnAsnHisPheAsnAlaLaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
396 |||||
602 AACACCACTTCAATGACGCTGTGGGTCCGATGTCACTCGATTTGATCTTGTGGGA 661
603 |||||
415 SerSerIleLeuValCysLeuProAsnGlyLeuTrpSerGlySerGluSerThrCysArg 434
416 |||||
662 AGCAGCATCATCTTATGTCTACCAATGGTTTGTGGTTCGGTTTACAGAGCTACTGCAGA 721
663 |||||
435 ValArgThrCysProHisLeuValGlnProLysHisGlyHisIleSerCysSerThrArg 454
436 |||||
722 GTAAGAACATGTCTCTCATCTCCGCCAGCCGAAACATGGGCCACATCAGCTGTTCTACAGG 781
723 |||||
455 GluMetLeuValCysLeuValAlaCysAspGluGlyValArgLeuGluGly 474
456 |||||
782 GAATGTATATAGACACATGTTTGTGTCTGTGATGAGGTTACAGAGCTAGAGGC 841
783 |||||
475 SerAspLysLeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysVal 494
476 |||||
842 AGTCATAAGCTTACTTGTCAAGGAAACAGCCAGTGGGATGGGCCAGAACCCCGTGTGTG 901
843 |||||
495 GluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCys 514
496 |||||
902 GAGCGCCACTGTITCCACCTTTTCAGATGGCCCAAGATGTCAATATCCCCCAACTGT 961
903 |||||
515 GlyLysGlnProAlaLysPheGlyThrIleCysValSerCysArgGlnGlyPheIle 534
516 |||||
962 GGCACAGCCAGCCCAAAATTTGGGACGATCTGCTATGTAAGTTGCGCCAGGGTTCAIT 1021
963 |||||
535 LeuSerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyVal 554
536 |||||
1022 TTATCTGGAGTCAAGCAAAATCTGAGATGTACCACTTCTCGAAAATGGAATGTGCGAGTT 1081
1023 |||||
555 GlnAlaAlaValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGlu 574
556 |||||
1082 CAGGAGCTGTGTGTAAGACAGTGGAGGCTCTCAATCAACTGTCTAAGGACATAGAG 1141
1083 |||||
575 AlaLysThrLeuGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLys 594
576 |||||
1142 GCTAAGACTCTGGAACAGCAGAGTCTGCAATGTTACCTGGCAGATTCACACAGCTAAA 1201
1143 |||||
595 AspAsnSerGlyGluLysValSerValHisValHisProAlaPheThrProProThrLeu 614
596 |||||
1202 GACAACTCTGGTGAAGGTGTCACTCCAGCTTCCAGCTTTCACCCCACTACCTT 1261
1203 |||||
615 PheProIleGlyAspValAlaIleValThrAlaThrAspLeuSerGlyAsnGlnAla 634
616 |||||

1262 TTCCCAATTGGAGATGTTGCTATCTGTATACACGGCAACTGACCTTATCGGGCAACAGGCC 1321
1263 |||||
635 SerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArg 654
636 |||||
1322 AGCTGCATTTTCCATATCAAGGTTATTTGATGCAGAACCACTGTTCATAGACTGGTGCAGA 1381
1323 |||||
655 SerProProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGln 674
656 |||||
1382 TCTCCACCTCCCGTCCAGGTTCTCGGAGAGGTATATGCCCCAGCTGGGATGAGCCCTCAG 1441
1383 |||||
675 PheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu 694
676 |||||
1442 TTTCTCAGACCACTCAGGGGCTGAATTTGGTCTATTACCAAGAGTCTATACCAAGGAGCTT 1501
1443 |||||
695 PheProGlnGlyGluThrIleValGlnThrAlaThrAspProSerGlyAsnAsnArg 714
696 |||||
1502 TTCCCTCAAGGGGAGACTATAGTACAGTATACGGCCACTGACCCCTCAGGCAATAACAGG 1561
1503 |||||
715 ThrCysAspIleHisIleValIleLysGlySerProCysGluIleProPheThrProVal 734
716 |||||
1562 ACATGTGATATCCATATTTGTATAAAGGTTCTCCCTGTGAAATTTCCATTCACACCTGTA 1621
1563 |||||
735 AsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeu 754
736 |||||
1622 AATGGGATTTTATATGCACTCCAGATATATCTGGAGTCACTGATTAACCTTGTCTTG 1681
1623 |||||
755 GluGlyValAspPheThrGluGlySerThrAspLysThrValAlaValGlyGluAspGly 774
756 |||||
1682 GAGGGCTATGATTTTCAAGAGGGTCTACTGACAGATTTATTTGTGCTTTATGAAGATGGC 1741
1683 |||||
775 ValTrpLysProThrThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsn 794
776 |||||
1742 GTCTGGAACCAACATATATACCTGAATGGCCAGACTGTGCCAAAAACGTTTTCGAAC 1801
1743 |||||
795 HisGlyPheLysSerPheGluMetPheValLysAlaAlaArgCysAspAspThrAspLeu 814
796 |||||
1802 CACGGTTTCAGTCTTGTGATGTTCTAAGAGAGCTGTTGTGATGACACAGATCTG 1861
1803 |||||
815 MetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMetValProSerPheCys 834
816 |||||
1862 ATGAAGAAGTTTCTCAAGCATTTTGAGACGACCTTGGGAAAAATGGTCCCATCATTTGT 1921
1863 |||||
835 SerAspAlaGluAspIleAspCysArgLeuGluAsnLeuThrLysLysThrCysLeu 854
836 |||||
1922 AGTGATGCAGAGACATTCGACCTGGAGAGAACCTGACCAAAAAATATTGCTTA 1981
1923 |||||
855 GluValAsnThrAspThrGluAsnGlyPheAlaIleGlyProGlyGlyTrpGlyAlaAla 874
856 |||||
1982 GAATATAATTATGACTATGAAATGGCTTTGCAATGGACAGGTGGCTGGGTGCAGCT 2041
1983 |||||
875 AsnArgLeuAspThrSerThrAspAspPheLeuAspThrValGlnGluThrAlaThrSer 894
876 |||||
2042 AATAGGCTGGATTTACTTCTTAAGATGATCTTCTGGACACTGTGCAAGAAACAGCCACAAGC 2101
2043 |||||
895 IleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspThrLysIle 914
896 |||||
2102 ATCGGCAATGCCAAGTCTCTCAGCGATTAAAGAGTGGCCCCATTTCTGACTATAAAT 2161
2103 |||||
915 LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeu 934
916 |||||
2162 AAGTTAATTTTACATCACAGCTAGTGTGCCATTACCCGATGAAGAAATGATACCCTT 2221
2163 |||||
935 GluTrpGluAsnGlnArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLys 954
936 |||||
2222 GAATGGGAAATCAGCAACGACTCTCTCAGACATTGGAAACTATCACAATAAATCTGAAA 2281
2223 |||||
955 ArgThrLeuAsnLysAspProMetThrSerPheGlnLeuAlaSerGluIleLeuIleAla 974
956 |||||
2282 AGGACTCTCAACAAAGACCCCATGTATTCCTTTTCAGCTGTCATCAGAAATATCTTATAGCC 2341
2283 |||||
975 AspSerAsnSerLeuGluThrLysLysAlaSerProPheCysArgProGlySerValLeu 994
976 |||||
2342 GACAGCAATTCATTAGGAACAAAAAGGCTTCCCTCTCTGACAGCCAGGCTCAGTCTG 2401
2343 |||||

```

QY 995 ArgGlyArgMetCysValAsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThr 1014
DB 2402 AGAGGCGGTATGTGTCAATTTGCCCTTTGGGAACCTATTATAATCTGGAAACATTTCCACC 2461
QY 1015 CysGluSerCysArgIleGlySerTyrGlnAspGluGlyGlnLeuGluCysValSerLeu 1034
DB 2462 TGTGAAGAGCTGCGGATCGGATCCCTATCAAGATGAAGAGGCAACTTGGAGTGAAGCTT 2521
QY 1035 CysProSerGlyMetTyrThrGluTyrIleHisSerArgAsnIleSerAspCysValAsa 1054
DB 2522 TGCCCTCTGGGATGTACAGGAATATATCCATTCAAGAAACATCTCTGATTGTAAAGCT 2581
QY 1055 GlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeu 1074
DB 2582 CAGTGTAAACAGGACCACTACTCATCGAGTGGACTTGTGAATCGTGTCCACTG 2641
QY 1075 GlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAsnThrSer 1094
DB 2642 GGCACCTTATCAGCCAAAATTTGGTTCCCGAGCTGCCCTCTCGTGTCCAGAAACACCTCA 2701
QY 1095 ThrValLysArgGlyValAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLys 1114
DB 2702 ACTGTGAAGAGGAGCGGTGAACATTTCTGCATGTGGAGTTCCTTGTCCAGAGGAAA 2761
QY 1115 PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrGlnProAsn 1134
DB 2762 TTCTCGCGTTCTGGGTTAAATGCGCTGTCAACCATGTCTCTCGTGAATATTATACCACTAAT 2821
QY 1135 AlaGlyLysAlaPheCysLeuAlaCysProPheTyrGlyThrThrProPheAlaGlySer 1154
DB 2822 CGAGGAAGGCGCTTCTGCTGCGCTGTCTCTTTATGGAACCTACCCCATTCCTCGTGTCC 2881
QY 1155 ArgSerIleThrGluCysSerSerPheSerThrPheSerAlaAlaGluSerVal 1174
DB 2882 AGATCCATCACAGATGTTCAAGTTTATAGTTCAACTTTCTCAGCGCGAGAGGAAAGTGTG 2941
QY 1175 ValProProAlaSerLeuGlyHisIleLysLysArgHisGluIleSerSerGlnValPhe 1194
DB 2942 GTGCCCCCTGCGCTCTCTTGGACATATTAAGAGGCGCATGAATCAGCAGTCAAGGTTTC 3001
QY 1195 HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGly 1214
DB 3002 CATGAATGCTCTTTAAACCTTGCCACATAGTAGGAACCTGCGACAACTTGGGCGTGTG 3061
QY 1215 TyrValCysLeuCysProLeuGlyTyrThrGlyLeuLysCysGluThrAspIleAspGlu 1234
DB 3062 TATGTTTGTCTCTGCTCCACTTGGATATACAGGCTTAAAGTGTGAAGACATCGATGAG 3121
QY 1235 CysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIle 1254
DB 3122 TGCAGCCCACTGCGCTTGCTCAACAATGAGGTTTGTAAAGACCTAGTTGGGGAATTCATT 3181
QY 1255 CysGluCysProSerGlyTyrThrGlyGlnArgCysGluGluAsnIleAsnGluCysSer 1274
DB 3182 TGTGAGTGGCCCATCAGGTTTACACAGCTACGCGGTGTGAAGAAATATATAATGAGTGTAGC 3241
QY 1275 SerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyrArgCysThr 1294
DB 3242 TCCAGTCTCTGTTTAAATAAAGGAATCTGTGTTGATGGTGTGGCTGCTATCTGTTGCACA 3301
QY 1295 CysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsn 1314
DB 3302 TGTGTGAAGAGGATTTGTATAGGCGCTGCAATTTGTGTGAACAGAGTCAATGAATGCCAGTCAAC 3361
QY 1315 ProCysLeuAsnAsnAlaValCysGluAspGlnValGlyGlyPheLeuCysValSerPro 1334
DB 3362 CCAATGCTTAAATATGACGCTCTGTGAGACCAAGTGTGGGGATTTCTTGTGCAATGCCCCA 3421
QY 1335 ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCys 1354
DB 3422 CCTGANTTTTGGGTACCGGATGTGGAAGAACGTCGATGAGTGTCTCAGTCAAGCCATGCG 3481

```

```

QY 1355 LysAsnGlyValaThrCysLysAspGlyAlaAsnSerPheArgCysLeuCysAlaAlaGly 1374
DB 3482 AAAAATGGAGCTACTCTGTAAAGACCGTGGCAATAGCTTTCAGATGCGTGTGTGACGTGGC 3541
QY 1375 PheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsn 1394
DB 3542 TTCACAGGATCACATCTGTGATTTGAACATCAATGAATGTCTAGTCTCAATCCATGTAGAAAT 3601
QY 1395 GlnAlaThrCysValAspGluLeuAsnSerTyrSerCysLysCysGlnProGlyPheSer 1414
DB 3602 CAGGCCACCTGTGTGATGAATTTAAATTCATACAGTGTGAATGTCTAGCCAGGATTTTCA 3661
QY 1415 GlyGlnArgCysGluThrGluClnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
DB 3662 GCGAAAAGGTGTGAACACAGACAGCTTACAGGCTTTAACTGGATTTTGAAGTTTCTGGC 3721
QY 1435 IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
DB 3722 ATCTATGATATGTATGTCTAGTTGGCATGTCTCCCATCTCTCATGCTCTTACCTGTACC 3781
QY 1455 PheTrpMetLysSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAsp 1474
DB 3782 TTCTGATGAATCTCTGACGACATGAACATGATGAACACCAATCTCTATGCACTGAT 3841
QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
DB 3842 AACGCGACGCAATATACCTTGTCTGCTGATATATACGGCTGGGTTCTTTATGTGAAT 3901
QY 1495 GlyArgGlyLysIleThrAsnCysProSerValAsnAspGlyValArgTrpHisIleAla 1514
DB 3902 GCGAGGAAAGATTAACAACTGCTCCTCGTGAATGATGGCAGATGGCATCATATTGCA 3961
QY 1515 IleThrTrpThrSerAlaAsnGlyIleTrpLysValTyrIleAspGlyLysLeuSerAsp 1534
DB 3962 ATCACTTGACAAAGTGCATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 4021
QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
DB 4022 GGTGTGTGTGCTCTCTGTGTTTGGCTTGGCCATACCTGCT 4060

```

RESULT 14

AK027870 5124 bp mRNA linear PRI 01-AUG-2002
 Locus Homo sapiens cDNA FLJ14964 fig. clone PLACE4000581, moderately
 similar to FIBROBLASTIN 1 PRECURSOR.

ACCESSION

AK027870.1 GI:14042858
 oligo capping: file (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Isoqai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
 Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
 Matsumura, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
 Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuno, Y. and Sasaki, N.
 NEDO human cDNA sequencing project

TITLE

Unpublished

JOURNAL

2 (bases 1 to 5124)

AUTHORS

Isoqai, T. and Otsuki, T.

JOURNAL

Submitted (10-MAY-2001) Takao Isoqai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kibarasu, Chiba 292-0812, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology; cDNA library construction,
 5'- & 3'-end one pass sequencing and clone selection: Helix
 Research Institute (supported by Japan key technology Center etc.)
 and Department of Virology, Institute of Medical Science,
 University of Tokyo.

FEATURES

Location/Qualifiers

source

1. 5124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE4000581"
 /tissue_type="placenta"
 /clone_lib="PLAC84"
 /note="cloning vector: pME18SFL3"
 119..4069
 /note="unmated protein product"
 /codon_start=1
 /protein_id="BAB55420.1"
 /db_xref="GI:14042859"
 /translation="WASTPKEEHCYLLHSFBEFEALARPALHEDLPSCFSFIODDMVHC
 SYLDEGDCDDRMGSCKGTHTGHPECICBKGTYGKGLQYECTACTCPSTYKPEGSPG
 GISSICPDENHTSPFGSTSPEDCVREGYRASGQTCLVHCPCALUPENGPFIQNT
 CNHFNACGVRCHPGLVGSIIILCLPNGLMSGLESYCRVETPHLRQPKRHHSIC
 STREMLYKTKTCLVACDEGTVLRGSKLTCQGNQMDPGRPCVVERHCSTFQMPKDVII
 SPHNGKOPAKFGTITCVSCROGFIILSGVKEMLCRTISGNVNVQVAAVCKDVEAPOI
 NCPDLIEAKTLEQDSNVTHQIPLTAKNGSEKVSVHVHPATPPVLPIDGVAVLYT
 ATDLGNQASCTIFHRIKVIDAEPVDMWCSPPPVQVSEKVAASWDEPQFSNAGNEL
 VITRSHNTQDLFPQGETIYQVYATDPSGNRTCDIHIVKSGSPCEIIPFVNGDFICT
 PONTGVNCTLCLEGDTFTEGTDKYCYAYEDGVWPKPTTTEWPCAKKRFPAHFGKS
 PMFYKAACDDTDLMKFSEAPETFLGKWPVFCSDARDIDWRLEENLTKKYCLLEYN
 VDYENGAIGPCGMAANRLDYSYDDLTVOETATISIGNAKSSRIKRSAPLSDYKIK
 LIPNTASVLPDERNDLWENQRLQTLLETINKLRLTNKOPWISFQASBILI
 ADSNSLGTAKSFPCKRPGSVLRGMVCNCPLGTYTNLEHFTCESKIGISTQREBQLE
 CKLSPGMYETIYHRSNTSDCAKQKQYTCGSLTSCPLGTYPQKFGSRSLSC
 TPTTVKRGANIISACQVPCPEGFVSRGLAPCHPCPRDYQPNAGAKFACIACPFYQ
 TTPFASGSIITFCSFSSTFSAABSVVPSLGHIKKRHISQVWFHCFNFCPNCHNS
 QCCQOLGRVYCLPLGYNLKACETDIDRCSPLCLNNGVCKDLNGEPICEKPSGVTV
 QRCEINNECSSPCLNGIICVDVAGRYCTCVKGFVGLHCETWENECQSNCLNNAV
 CEDQVGFELCKPFLGTRCNGNDECLSPCKNGATCKDGNFRCLCAAGFTGSH
 CELNINBQSNCRQATCDLNSYCKCPFGSGRCETQSTQFNLDFFVSGIYG
 YVMLVPLPSLHALACTFMKSSDDMYTGPISYAVDNGSDNTLLITDYNQWLVYNG
 REKINCPSPVNDGRWHIIALTWTSANGIKWYIDGKLSDGAGLSVGLPIPGMF"

CDS

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 5124
 Score: 7537.00 Matches: 1347
 Percent Similarity: 99.63% Conservative: 1
 Best Local Similarity: 99.56% Mismatches: 5
 Query Match: 37.74% Indels: 0
 DB: 9 Gaps: 0

US-09-977-053-4 (1-3571) x AK027870 (1-5124)

QY 195 ThrAspGlyTyrSerAsnGlyValAspProArgProIleAlaSerLeuArgAspSer 214
 DB 2 ACTGATGGATATTCATATGGGGAGACCCCTAGACCAATTCAGCGCTCACTCGAGATTCA 61
 QY 215 GlyValGluIlePheThrPheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMet 234
 DB 62 GGAGTGGAGATCTTCACATTTTGGCATATGCAAGGGAACATTCGAGAGCTGAATGACATG 121
 QY 235 AlaSerThrProLysGluIleHisCysTyrLeuLeuHisSerPheGluGluPheGluAla 254
 DB 122 GCTTCCACCCCAAGGAGGAGCACTGTATTACCTGCTACACAGATTTTGAAGAAATTTGAGGCT 181
 QY 255 LeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMet 274
 DB 182 TTAGCTCCCGGGCATGATGATGAGATCTACCTTTCTGGGAGTTTATTTCAGATGATATG 241
 QY 275 ValHisCysSerTyrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCys 294
 DB 242 GTCCACTGCTCATATCTTTGTCATGAAGCAAGGAGCTGTGTGACCGCAATGGAGCTGC 301
 QY 295 LysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLys 314
 DB 302 NAATGTGGGACACACACAGGCCATTTTGAGTGCATCTGTGAAAAGGGGGTATTACGGGAAA 361
 QY 315 GlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySerPro 334

DB 362 GGTCTGCAGTATGATGACAGCTTGGCCATCGGGGACATACAACTCGAGGCTCACCA 421
 QY 335 GlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProProGlySer 354
 DB 422 GGAGGAATCAGCAGTGTGCATTCCTCTGATGAAATCACACTCTCCACCTCGAAGC 481
 QY 355 ThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGlu 374
 DB 482 ACATCCCTCGAAGACTGTGTCTGAGAGAGGGATACAGGGCATCTGCCAGACTGTGAA 541
 QY 375 LeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCys 394
 DB 542 CTGTGTCACCTGCTGCCCTGAAGCTCTCCGAAATGGTTACTTTATCCAAACACTTGC 601
 QY 395 AsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGly 414
 DB 602 AACACCACTTCAATGTCAGCTGTGGGGTCCGATGTCACTCGATTTGATCTTGTGGGA 661
 QY 415 SerSerIleIleLeuCysLeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArg 434
 DB 662 AGCAGCATCATCTTATGTCTTACCAATGGTTTGTGGTCCGGTTTAGAGAGCTACTGAGA 721
 QY 435 ValArgThrCysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThrArg 454
 DB 722 GTTAGAACATGCTCTCATCTCCGCCAGCGAACAATGSCCACCACATCAGCTGTTCTACAAG 781
 QY 455 GluMetLeuTyrLysThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGly 474
 DB 782 GAAATGTTATATAAGACAACATCTTTGTTGTCCTGTGATGAAGGTACAGACTAGAAAGC 841
 QY 475 SerAspLysLeuThrCysGlnGlyAsnSerGlnTyrAspGlyProGluProAspCysVal 494
 DB 842 AGTGATAAGCTTACTTGTCAAGGAAACAGCCAGTGGGTGGGCGAAGCCCGGTGTGTG 901
 QY 495 GluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAsnCys 514
 DB 902 GAGGCGCACTGTTCCACCTTTCAGATGCCCAAGATGTTCATATATCCCCCACTGT 961
 QY 515 GlyLysGlnProAlaLysPheGlyThrIleCysTyrValSerCysArgGlnGlyPheIle 534
 DB 962 GGCAAGCAGCCAGCCAAATTTGGGACGATCTGTATGTAAAGTTGCCGCCAAGGGTTCA 1021
 QY 535 LeuSerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTyrAsnValGlyVal 554
 DB 1022 TTATCTGGAGTCAAGAAATGCTGAGATGTACCACTTCTGGAATAATGGAATGTCCGAGTT 1081
 QY 555 GlnAlaAlaValCysLysAspValGluAlaProGlnIleAsnCysProLysAspIleGlu 574
 DB 1082 CAGGCGAGCTGTGTGTAAGACGTGGAGGCTCTCTCAATCAACTGTCCTAAGGACATAGAG 1141
 QY 575 AlaLysThrLeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLys 594
 DB 1142 GCTAAGACTCTGAAACAGCAAGATTCTGCCAATGTATTACCTGGCAGATTTCAACAGCTAAA 1201
 QY 595 AspAsnSerGlyLysValSerValHisValHisProAlaPheThrProProTyrLeu 614
 DB 1202 GAACTCTGGTGAAGGTGTGCTCCAGCTTCATCCAGCTTTCACCCCACTTACCTT 1261
 QY 615 PheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAla 634
 DB 1262 TTCCCAATTTGGAGATGTGTCTATCGTATACAGGCAACTGACCTATCCGCAACACAGGCC 1321
 QY 635 SerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTyrCysArg 654
 DB 1322 AGCTGCAATTTCCATATCAAGGTATTGTATGTCAGAAACCACTCTGTCTAGACTGTGTGAGA 1381
 QY 655 SerProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGln 674
 DB 1382 TCTCCACCTCCCGTCAGGTCTCGGAGAGGTATACGCCCAAGCTGGGATGAGCCTCAG 1441
 QY 675 PheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeu 694

Db	1442	TTCTCAGCAAACTCAGGGCTGAAATTTGGTCAATTACCAAGAGTCAATACCAAGGAGACCTT	1501
Qy	695	PheProGlnGlyGluThrIleValGlnIlyrThrAlaThrAspProSerGlyAsnAsnArg	714
Db	1502	TTCCCTCAAGGGAGACTATAGTACAGTATACGGCCACTGACCCCTCAGGCAATAACAGG	1561
Qy	715	ThrCysAspIleHisIleValIleIleGlySerProCysGluIleProPheThrProVal	734
Db	1562	ACATGTGATATCCATATTGTTCATAAAAGGTTCTCCCTGTGAAATTCATATTCACACCTGTA	1621
Qy	735	AsnGlyAspPheIleCysThrProCysAsnThrGlyValAsnCysThrLeuThrCysLeu	754
Db	1622	AATGGGGAATTTATGTGACCTCCAGATAATACTGGAGTCAACTGTACATTAACCTTGTCTG	1681
Qy	755	GluGlyTyraSpPheThrGluGlySerThrAspIlySerTyrCysAlaTyrgluAspGly	774
Db	1682	GAGGGCTATGATTTACAGAAGGGTCTACTGACAAGTATTATTGTGCTTATGAGATGGC	1741
Qy	775	ValTrpIysProThrTyThrThrGluTrpProAspCysAlaIlysIysAlaGpheAlaAsn	794
Db	1742	GTCTGGAAACCACATATACCCTGAAATGGCCAGACTGTGCCAAAAAACGTTTGTCAAAC	1801
Qy	795	HisGlyPheIysSerPheGluMetPheTyrylsalalaargCysAspAspThrAspLeu	814
Db	1802	CACGGGTTCAAGTCCCTTTGAGATGTTCTACAAACAGCTGCTGTTGTATGACACACAGATCTG	1861
Qy	815	MetIlysIysPheSerGluAlaPheGluThrThrLeuGlyIysMetValProSerPheCys	834
Db	1862	ATGAAGAAGTTTCTCGAAGCANTTTGAGACGACCTTGGGNAANATGGTCCCATCATTTTGT	1921
Qy	835	SerAspAlaGluAspIleAspCysArgIleGluGluAsnLeuThrIlysTyryCysLeu	854
Db	1922	AGTGATGACAGAGGACATTGACTCGAGACTGGAGGAGAACCTTGACCAAAAAATATTGGCTA	1981
Qy	855	GluTyraSerTyraSpTyrgluAsnGlyPheAlaIleGlyProGlyGlyTrpGlyAlaAla	874
Db	1982	GAATATATAATTATGACTATGAAAAATGGCTTTGCAATTGGACAGGTGGCTCGGGGTGAGCT	2041
Qy	875	AsnArgLeuAspTyrySerTyraSpAspPheLeuAspThrValGlnGluThrAlaThrSer	894
Db	2042	AATAGGCTGGATTTACTCTTACATGACTTCTTGACACTGTGCUAGAAACAGCCACAAAGC	2101
Qy	895	IleGlyAsnAlaIysSerArgIleIlysaArgSerAlaProLeuSerAspTyryLysIle	914
Db	2102	ATCGCAATGCCAAGTCTCAAGGNTAAAGAAGTGCCTTATCTGACTATAAAATTT	2161
Qy	915	LysLeuIlePheAsnIleThrAlaSerValProLeuProAspGluArgAsnAspThrLeu	934
Db	2162	AAAGTTAATTTTTTAAACATCACAGCTAGTGTGTCATTTACCCGATGAAGAAATGATACCCCTT	2221
Qy	935	GluTrpGluAsnGlnArgLeuGlnThrLeuGluThrIleThrAsnIlysLeuIys	954
Db	2222	GAATGGGNAATCAGCAACGACTCTTTCAGACATTGGAAACTATCAAAAATAAATCTGAAA	2281
Qy	955	ArgThrLeuAsnIysAspProMetTyrySerPheGlnLeuAlaSerGluIleLeuIleAla	974
Db	2282	AGGACTCTCAACAAAGACCCCATGTATTCTTTTCAGCTTGCATCAGAAATACTTATAGCC	2341
Qy	975	AspSerAsnSerLeuGluThrLysGlyAlaSerProPheCysArgProGlySerValLeu	994
Db	2342	GACAGCAATTCATTTAGGNAACAAAAAAGGCTTCCCTTCTGCGACAGCAGGCTCAGGTGCTG	2401
Qy	995	ArgGlyArgMetCysValAsnCysProLeuGlyThrTyryTrpAsnLeuGluHisPheThr	1014
Db	2402	AGAGGGGTATGTGTGTCAATTGGACCTTTGGAAACCTATTATATCTTGGAAACATTTTCACC	2461
Qy	1015	CysGluSerCysArgIleGlySerTyrglnAspGluGlyGlnLeuGluCysIysLeu	1034
Db	2462	TGTCAAAGCTGCCGGATCGGATCTCATCAAGATCAAGAGGCGCAACTTGATGTCAAGCTT	2521
Qy	1035	CysProSerGlyMetTyThrGluTyryIleHisSerArgAsnIleSerAspCysIysAla	1054
Db	2522	TGCCCTCTGGGATGTACAGGAATATCCATTTCAAGAAACATCTCTGATTTGTAAGCT	2581

QY	1055	GlnCysLysGlnGlyThrTyrserTyrsGlyLeuGluThrCysGluSerCysProLeu	1074
DB	2582	CAGTGTAAACAAGGCACCTACTCATGTGAGTGGAGCTTGAGCAATCGTGTCCACTG	2641
QY	1075	GlyThrTyrglnProLysPheGlySerArgSerCysLeuSerCysProGluLeuAsnThrSer	1094
DB	2642	GGCACTTATCAGCCCAAAATTGGTTCCGGAGAGCTGCCTCTCGTGTCCAGAAACACCTCA	2701
QY	1095	ThrValLysArgGlyAlaValAsnIleSerAlaCysGlyValProCysProGluGlyLys	1114
DB	2702	ACTGTGAAAGAGAGCGCGTGAACATTTCTGCATGTGAGTTCTCTTGTCAGGAAGGAAAA	2761
QY	1115	PheSerArgSerGlyLeuMetProCysHisProCysProArgAspTyrTyrglnProAsn	1134
DB	2762	TTCTCGGCTTCGGGTAAATGCCCTGTCAACCCATGCTCGTACATTAACCAACCTAAT	2821
QY	1135	AlaGlyLysAlaPheCysLeuAlaCysProPheTyrglyThrThrProPheAlaGlySer	1154
DB	2822	GCAGGGAAGGCCTTCTGCCTGGCTGTCCCTTTATGGAACACTACCCCATTCGCTGGTTC	2881
QY	1155	ArgSerIleThrGluCysSerSerPheSerSerThrPheSerAlaAlaGluGluSerVal	1174
DB	2882	AGATCCATCAACAGAAATTCACAGTTTATGTTTCAACTTCTCAGCGGCAGAGGAAGTGTG	2941
QY	1175	ValProAlaSerLeuGlyHisIleLysValArgHisGluIleSerSerGlnValPhe	1194
DB	2942	GTGCCCTCGCTCTCTTGACATATTAAAGAGGCGTGAATCAGCACTCAGTTTTC	3001
QY	1195	HisGluCysPhePheAsnProCysHisAsnSerGlyThrCysGlnGlnLeuGlyArgGly	1214
DB	3002	CATCAATGCTTCTTTAAACCTTGCACAAATAGTGAACCTGCCAGCAACTTGGCGGTGT	3061
QY	1215	TyrValCysLeuCysProLeuGlyTyThrGlyLeuLysCysGluThrAspIleAspGlu	1234
DB	3062	TATGTTGTCTGTGTCACTTGGATATACAGGCTTAAAGTGTGAACACAGATCGATGAG	3121
QY	1235	CysSerProLeuProCysLeuAsnAsnGlyValCysLysAspLeuValGlyGluPheIle	1254
DB	3122	TGCAGCCCACTGCCCTCAACAATGAGTGTGTAAAGACCTAGTTGGGGAATTCATT	3181
QY	1255	CysGluCysProSerGlyTyThrGlyLysArgCysGluLysIleAsnGluCysSer	1274
DB	3182	TGTGAGTGCCTCATCAGGTATACAGGTGCACGGTGTGAAGAAATATAATGATGATGTAGC	3241
QY	1275	SerSerProCysLeuAsnLysGlyIleCysValAspGlyValAlaGlyTyArgCysThr	1294
DB	3242	TCCAGTCTCTGTTTAAATAAGGAATCTGTGTATGGTGTGCTGCTATCTGTGCACA	3301
QY	1295	CysValLysGlyPheValGlyLeuHisCysGluThrGluValAsnGluCysGlnSerAsn	1314
DB	3302	TGTGTGAAGAGATTTGTAGCGCTGCATTTGTGAACACAGAACTCAATGAATGCCAGTCAAA	3361
QY	1315	ProCysLeuAsnAsnAlaValCysGluAspGluValGlyClyPheLeuCysLysCysPro	1334
DB	3362	CCATGCTTAAATATGACAGTCTGTGAAGACCAAGTGTGGGGATCTTTGTGCCAATGCCCA	3421
QY	1335	ProGlyPheLeuGlyThrArgCysGlyLysAsnValAspGluCysLeuSerGlnProCys	1354
DB	3422	CCTGGATTTTTGGGTACCCGATGTGGAAGAACAGTTCGATGAGTGTCTCAGTCAGCCATGC	3481
QY	1355	LysAsnGlyAlaThrCysLysAspGlyValaAsnSerPheArgCysIleuCysAlaAlaGly	1374
DB	3482	AAAAATGGNGCTACCTGTGTAAAGACGGTGCCTTCAAGTTTCAGATGGCTGTGTGCAGCTGGC	3541
QY	1375	PheThrGlySerHisCysGluLeuAsnIleAsnGluCysGlnSerAsnProCysArgAsn	1394
DB	3542	TTACAGAGTACACACTGTGTGATTTGAACATCAATGAATGTCTCAGTCTAATCCATGTAGAAAT	3601
QY	1395	GlnAlaThrCysValAspGluLeuAsnSerTyrsCysLysCysGlnProGlyPheSer	1414
DB	3602	CAGGCCACCTGTGTGTGAATGAATTAATTTATACATCAGTTGTAAATGTGCAGCAGATTTTCA	3661

QY 1415 GlyGlnArgCysGluThrGluGlnSerThrGlyPheAsnLeuAspPheGluValSerGly 1434
 |||::|||
 Db 3662 GCGAAAGGTGTGAACACAGAGTCTACAGGCTTTAACTGATTTGAAGTTTCTGGC 3721
 QY 1435 IleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThrCysThr 1454
 |||::|||
 Db 3722 ATCTATGATATGTCATGCTAGTTGGCATGCTCCCATCTCTCCATCTCTTAACCTGTACC 3781
 QY 1455 PheTyrMetLeuSerSerAspAspMetAsnTyrGlyThrProIleSerTyrAlaValAsp 1474
 |||::|||
 Db 3782 TTCTGATCAAAATCTCTGACGACATGAATATGGAACCACTCTCTATGCAAGTTGAT 3841
 QY 1475 AsnGlySerAspAsnThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyrValAsn 1494
 |||::|||
 Db 3842 AACGGCAGGACAAATACCTTCTCTGCTGACTGATTAACAGCTGGGTCTTTATGATGAAT 3901
 QY 1495 GlyArgGluLysIleThrAsnCysProSerValAsnAspGlyArgTyrHisIleAla 1514
 |||::|||
 Db 3902 GGCAGGAAAGATAACAACCTGCTCCTCGTGAATGATGGCAGATGCATCATATTGCA 3961
 QY 1515 IleThrTyrThrSerAlaIleThrGlyValTyrIleAspGlyLysLeuSerAsp 1534
 |||::|||
 Db 3962 ATCACTTGGACAGTGCCTGCAATGTCATCTGAAAGTCTATATGATGCGAAATATCTGAC 4021
 QY 1535 GlyGlyAlaGlyLeuSerValGlyLeuProIleProGly 1547
 |||::|||
 Db 4022 GGTGGTGTGGCTCTCTGTGTGTTTGCCCATACCTGGT 4060

RESULT 15

LOCUS HSM806219 4385 bp mRNA linear PRI 17-JUN-2003
 DEFINITION Homo sapiens mRNA; cDNA DKFZp686L15151 (from clone DKFZp686L15151).
 ACCESSION BX538049
 VERSION BX538049.1 GI:31874164

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 4385)
 Blocher, H., Boecker, M., Mewes, H.W., Weil, B., Amlid, C., Osanger, A.,
 Fobo, G., Han, M. and Wiemann, S.

REFERENCE

AUTHORS

Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuherberg, GERMANY

TITLE

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by GSF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.

This clone (DKFZp686L15151) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
 information about the clone and the sequencing project is available
 at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..4385
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /map="9q31.2-32"
 /clone="DKFZp686L15151"
 /tissue_type="human colon endothel primary cell culture"
 /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
 DH10B; sites SfiIA + SfiIB"
 /dev_stage="adult"
 1..4385
 /genes="DKFZp686L15151"
 <1..3972
 /genes="DKFZp686L15151"
 /notes="hypothetical protein, N-terminus truncated"
 /codon_start=1
 /product="hypothetical protein"

gene

CDS

/protein_id="CAD97988.1"
 /db_xref="GI:31874165"
 /translation="RHHSESLPMKVPDLCKPKPPPTQNGPMKGFVGVSKVQPPCNE
 GYELVGGSWTCKQKNNKSNPKMPKPCPEPPLLENQLVKRLTEVGVVTFCK
 RGHVLQPSVLKCLPSQWNDSFPVCKI VLCTPPLISFGVPI PSSALHFGSTVYKSC
 KGFGLRGNSTTLCPDGTWSSPCKVPCVPCPBPI NGI IDVGLAYLSTALYTC
 KGFELRGNSTTLCPDGTWSSPCKVPCVPCPBPI NGI IDVGLAYLSTALYTC
 NGRFLRGNSTTLCPDGTWSSPCKVPCVPCPBPI NGI IDVGLAYLSTALYTC
 PFGVAGHAMWTCESSHSIPTCHPIDCGLPHIDFGDCTKLDKDDQGYFQEDDM
 MEVYVTPPHYHLGAVAKWENTKESPATHNSNLYGTNWSYTCNPGYELGNPVL
 CQEDGTWNSAFSCISIEDLPATNGFLRPTTETSMGSAVQSKCPGHILAGDLRL
 CLENNSGASPRCEAISCKKNPVNNGSIKGSNTYTLSTLYTECDPGYVLNGTERT
 CDDKWDDEBICI PVDSSPPVSANGVGRDEYTFKEIETVTCNCGFLLRGRSRV
 CLANGSWATPDVPCVRCATPOLANGVTGLDYGKMEVTFCHGEGYILHGAOKLT
 CDSGDNWDIIEFLCKPNCGPEDLAHPNGFSFHGHIOYQCPSYKLGHSRR
 CUSGNSGSSPCRCSTPVIEYGVNGTDFDCGAARIQCFKFGKLGJLSEITC
 RADGWSGFPCHCSTCSLPMINAFISETSSWKENVITSCRSYVITOGSSDLIC
 TRKWSQVYPCVPCPLSCGSPSVANAVATGEATYSEVKLRLEGYTMDDTDTFT
 CKDGRWPERISCSPKCPLESI THILVHGDDFSVNRQVSVSCARGYTFEGVNI SV
 COLDGTWSPSPDES CSPVSCGPSEPHGVVGVKYPETSTI IYQCEPGYELRGR
 RVCQNRQSGVAICKETRCETPLFLANGKADIENTTGPNNVYSCNRYSLRGPSE
 AHCNTGTVSHVPVLPCKPNPFPVPI PENALISEKEFYVDQNVSI ICKRBFLLQGHG
 IITCNPDETWTSAKCKIKSCGPAPHVENAIJARGVHYQYGDIMITYCYGMLGFL
 RSVCLENGTTWTPPI CRAVCFPCQNGGICQRPNACSCPRGMGRLECEBPTCIPCLN
 GGRCAVAPYQDCDPPGWTGSRCHTAVCQSPCLNGCKVRPNRCHCLSSWTCHNCSGKR
 TGP"
 4205..4210
 /genes="DKFZp686L15151"
 4229
 /genes="DKFZp686L15151"
 polyA_signal
 polyA_site
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 4385
 Score: 7486.00 Matches: 1319
 Percent Similarity: 99.77% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 3
 Query Match: 37.48% Indels: 0
 DB: 9 Gaps: 0
 US-09-977-053-4 (1-3571) x HSM806219 (1-4385)

QY 2249 ArgHisTyrHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysProPro 2268
 |||::|||
 Db 1 CGCACCTGGCACAGTGAATCCCTCTGATGTGTGTTCTCTGACTGTGGAAAACCTCCC 60
 QY 2269 ProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPhe 2288
 |||::|||
 Db 61 CCGATCCAGATGGCTTCATGAAAGAGAGAAAACCTTTGAAGTAGGCTCCAGGTTTCAGTTT 120
 QY 2289 PheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTyrThrCysGlnLysSerGly 2308
 |||::|||
 Db 121 TTCTGTAATAGAGGTATGAGCTTGTGTGGCAGTTCTTGGACATGTCCAGAAATCTGGC 180
 QY 2309 LysTyrAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProProLeu 2328
 |||::|||
 Db 181 AAATGGATATAGAGTCAATCCAAAGTCATGCTCCCAAGTCCCAAGCCGCCCCCTC 240
 QY 2329 LeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSer 2348
 |||::|||
 Db 241 TTGAAAACACAGCTAGTATTATAAGGAGTTGACCAACCCAGGTTAGGAGTTGTGACATTTTC 300
 QY 2349 CysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGln 2368
 |||::|||
 Db 301 TGTAAAGAGAGGAGTGTCTGCAAGGCCCTCTGTCTCTGAAATGCTTGCCTCATCCAGCAA 360
 QY 2369 TrpAsnAspSerPheProValCysLysValIleValLeuCysThrProProLeuIleSer 2388
 |||::|||
 Db 361 TGGATGACTCTTCTCTGTTTGTAGATGTTCTTTGATCCCACTCCCTCAATTATTC 420
 QY 2389 PheGlyValProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCys 2408
 |||::|||
 Db 421 TTGTGTGTCCTGT 480

2409 ValGlyGlyPhePheLeuArgGlyAenSerThrThrLeuCysGlnProAspGlyThrTrp 2428
481 GTAGTGGGCTTTCTTAAGAGGAAATCTACCACTCTGCGAACCTGATGGCACCTGG 540
2429 SerSerProLeuProGluCysValProValGluCysProGlnProGluGluLeuProhan 2448
541 AGCTCTCCACTGCCAGATGTGTTCCAGTAGAATGTCCCAACCTCGAGAGAAATCCCAAT 600
2449 GlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysPro 2468
601 GGAATCATGTGTCAGAGGCTTGGCTATCTCAGCAGAGCTCTCTATACCTGCGAGCCA 660
2469 GlyPheGluLeuValGlyAenThrThrThrThrLeuCysGlyGluAenGlyHisTrpLeuGly 2488
661 GGCCTTTGAATTTGGTGGGAAATACTACCACTCTTTGGAGAGAAATGGTCACTGGCTTGA 720
2489 GlyLysProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAenGlyLys 2508
721 GGAATAACCAACATGTAAAGCCATTGAGTGCCTGGAACCCAGAGAGATTTTGAATGGCAAA 780
2509 PheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAenArgGlyPhe 2528
781 TTCTCTTACCGGACTACACTATGACAGACCGTTACTACTCTTGCACCGAGGCTTT 840
2529 ArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAla 2548
841 CGGCTCGAAGGTCCAGTGCCTTGACCTCTTTAGAGACAGGTGATGGGATGTAGATGCC 900
2549 ProSerCysAenAlaIleHisCysAspSerProGlnProIleGluAenGlyPheValGlu 2568
901 CCATCTTGAAATGCCATCCACTGTGATTCCTCCCAACCCAGAGAGATTTTGTAGAA 960
2569 GlyAlaAspTyrSerTyrGlyAlaIleIleTyrSerCysPheProGlyPheGlnVal 2588
961 GGTGCAGATTACAGTATGTTGCTCCATATCATCTACAGTTGCTTCCCTGGGTTTCAGGTG 1020
2589 AlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerIleProThrCys 2608
1021 GCTGTCTATGCTCATGTCAGACCTGTGAAGAGTCAGGATGGTCAAGTTCCATCCCAACATGT 1080
2609 MetProIleAspCysGlyLeuProHisIleAspPheGlyAspCysThrLysLeuLys 2628
1081 ATGCCAATAGACTGTGGCTCCCTCTCATATAGATTTTGGAGACTGTACTAAACTCAAA 1140
2629 AspAspGlnGlyTyrPheGluGlnGluAspMetMetGluValProTyrValThrPro 2648
1141 CATGACCAAGGNTATTTTGAGCAAGAGACATGATGGAGTTCCATATGTGACTCT 1200
2649 HisProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAenThrLysGluSerPro 2668
1201 CACCTCTCTTATCATTTGGAGCAGTGGCTAAACCTCGGAGAAATACAAAGAGTCTCT 1260
2669 AlaThrHisSerSerAspPheLeuTyrGlyThrMetValSerTyrThrCysAenProGly 2688
1261 GCTACACATTCATCAAACTTCTGTATGTGATGATCCATGGTTTCATACCTGTATATCAGA 1320
2689 TyrGluLeuLeuGlyAenProValLeuIleCysGlnGluAspGlyThrTrpAenGlySer 2708
1321 TATGAACCTCTGGGNAACCTGTGTGATCTCCAGGAGATGGAACTTGAATGGCAGT 1380
2709 AlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAenGlyPheLeu 2728
1381 GCACCATCTCGCATTTCAATTTGAATGTGACTTGGCTTACCTGCTCCCTGGAATGGCTTTTG 1440
2729 ArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyHisIle 2748
1441 CTTTTTACAGAGACTAGCATGGGAAGTCTGTGTCAGATATAGCTGTAAACCTGGACACAT 1500
2749 LeuAlaGlySerAspLeuArgLeuCysLeuGluAenArgLysTrpSerGlyAlaSerPro 2768
1501 CTAGCAGGCTCTGACTTAAGGCTTTGTCTAGAGAAATAGAAAGTGGAGTGGTGCCTCCCA 1560

2769 ArgCysGluAlaIleSerCysLysLysProAenProValMetAenGlySerIleLysGly 2788
1561 CGCTGTGAAGCCATTTTCATGCAAAAGCCAAATCCAGTCAATGATGATCCATCAAGA 1620
2789 SerAenTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAen 2808
1621 AGCAACTACATACCTGAGCACGTTGTACTATGAGTGTGACCCCGATATATGTCTGAAT 1680
2809 GlyThrGluArgArgThrCysGlnAspAspLysAenTrpAspGluAspGluProfileCys 2828
1681 GGCACCTGAGAGAGAGAAATGCCAGGATGACAAAATCGGATGAGGATGAGGCCATTTGC 1740
2829 IleProValAspCysSerSerProProValSerAlaAenGlyGlnValArgGlyAspGlu 2848
1741 ATTTCTGTGGACTGCGATTTCACCCCGAGTCTCAGCCAAATGGCCAGGTGAGAGAGAGAG 1800
2849 TyrThrPheGlnLysGluIleGluTyrThrCysAenGluGlyPheLeuLeuGluGlyAla 2868
1801 TACATTTCCAAAAGAGATTGATACACTTGCATATGAAGGTTCTTGTCTTGAAGGAGCC 1860
2869 ArgSerArgValCysLeuAlaAenGlySerTrpSerGlyAlaThrProAspCysValPro 2888
1861 AGGAGTGGGTTTGTCTTGGCAATGGAAATTTGGAGTGGAGCCACTCCCGACTGTGTGCCT 1920
2889 ValArgCysAlaThrProProGlnLeuAlaAenGlyValThrGluGlyLeuAspTyrGly 2908
1921 GTCAGATGTGCCACCCGCCCACTGGCCCAATGGGGTGACGGAGGCCCTGGACTATGGC 1980
2909 PheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProLys 2928
1981 TTCATGAGGAGATTAACATTCCACTGTCAAGAGGCTACATCTTGCACCGTCTCTCCAAA 2040
2929 LeuThrCysGlnSerAspGlyAenTrpAspAlaGluIleProLeuCysLysProValAen 2948
2041 CTCACCTGTCACTGAGTGGCAACTGGGATGCGAGATTCCTCTCTGTAAACCCAGTCAAC 2100
2949 CysGlyProProGluAenLeuAlaHisGlyPheProAenGlyPheSerPheIleHisGly 2968
2101 TGTGACCTCTCTGAAGATCTTGGCCCATGGTTTCCCTTAATGGTTTTTCTTTTATCATGG 2160
2969 GlyHisIleGlnTyrGlnCysPheProGlyTyrLysLeuHisGlyAenSerSerArgArg 2988
2161 GGCCATATACAGTATCAGTGTCTTCTCTGTTATAGCTCCATGGAATTCATCAAGAAG 2220
2989 CysLeuSerAenGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSer 3008
2221 TGCCTCTCCAAATGGCTCTCGAGTGGCAGCTCACCTTCTCTGCTGCTTGCAGATGTTC 2280
3009 ThrProValIleGluTyrGlyThrValAenGlyThrAspPheAspCysGlyLysAlaAla 3028
2281 ACACCAAGTAATGAATATGGAACTGTCAATGGGACAGATTTTGTACTGTGGAAGGCGAGCC 2340
3029 ArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAla 3048
2341 CGGATTCAGTGTCTTCAAGGCTTCAAGCTCTTAGGACTTTCTGAAATCACCTGTGAAGCC 2400
3049 AspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuPro 3068
2401 GATGGCCAGTGGAGCTCTGGGTTCCCCACTGTGAACACACTTCTTGTGGTTCTCTTCCA 2460
3069 MetIleProAenAlaPheIleSerGluThrSerSerTrpLysGluAenValIleThrTyr 3088
2461 ATGATACCAATATCGGTTTCATCTGAGGAGCCAGCTCTTGGAGGAGAAATGTGATTAACCTTAC 2520
3089 SerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGly 3108
2521 AGTGCAGGCTCTGGATATGTATACAGGCACTCAGATCTGATTTGTATACAGAGAAAGG 2580
3109 ValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerVal 3128
2581 GTATGAGGACGAGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTGCCCAACCGTCTGTC 2640
3129 AlaAenAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCys 3148

```

Db      2641  |||||GCCATGTCAGTGGCACTGGAGAGGCAACACCTATGAAAGTGAAGTGAACCTCAGATGT 2700
Qy      3149  |||||LeuGluGlyTyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyArg 3168
Db      2701  |||||CTGGAAGTTATACGATGATACAGATACAGATACATATTCACCTGTGACAGAAATGGTGGC 2760
Qy      3169  |||||TTPheProGluArgIleSerCysSerProLysCysProLeuProGluAsnIleThr 3188
Db      2761  |||||TGGTTCCTCGAGAGAAATCTCTCGAGTCCCTAATAAATGTCTCTCCCGGAAGCATAAACA 2820
Qy      3189  |||||HisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAla 3208
Db      2821  |||||CATATATCTGTACATGGGAGCATTTCTGATGTGATAGGCAAGTTCTGTGTCTATGTGCA 2880
Qy      3209  |||||GluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGlu 3228
Db      2881  |||||GAAGGGTATACCTTTGAGGGAGTTAAATATCATATCATGTATGTCTGAGCTTGATGAACTGGGAG 2940
Qy      3229  |||||ProProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGlu 3248
Db      2941  |||||CCACCAATCTCCGATGAATCTGAGTCCAGTTCTTGTGGGAACCTGAAGTCCAGAA 3000
Qy      3249  |||||HisGlyPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGlu 3268
Db      3001  |||||CATGGATTTGTGTGGTGGCAGTAATAACCTTTGAAAGCAATATTATTCAGTGTGAG 3060
Qy      3269  |||||ProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSer 3288
Db      3061  |||||CCTGGCTATGAATAGAGGGGAAACAGGAGCTGTCTGCCAGGAGACACAGAGTGGAGT 3120
Qy      3289  |||||GlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGly 3308
Db      3121  |||||GGAGGGGTGGCAATATGCAAAAGACAGAGTGTGAATCTCCATCTGAATTTCTCAATGGG 3180
Qy      3309  |||||LysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGly 3328
Db      3181  |||||AAGCTGTACATTTGAACACAGACAGCTGGACCCACGTTGTATATCTCTGCACAGAGGC 3240
Qy      3329  |||||TyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisPro 3348
Db      3241  |||||TACAGTCTTGAAGGGCCATCTGAGGCACACTGCACAGAAAATGAACTGGAGCCACCCA 3300
Qy      3349  |||||ValProLeuCysLysProAsnProCysProValProPheValIleProGluAsnAlaLeu 3368
Db      3301  |||||GTCCCTCTCTGCAAAACCAATTCATGCTCTTGTCTTGTGATTTCCCGAGAATGCTCTG 3360
Qy      3369  |||||LeuSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPhe 3388
Db      3361  |||||CTGTCTGAAAAGAGTTTATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420
Qy      3389  |||||LeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSer 3408
Db      3421  |||||CTCTGCGAGGCCACGGCATCATTTACCTGCAACCCCGACGAGCTGGACACAGACAAGC 3480
Qy      3409  |||||AlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArg 3428
Db      3481  |||||GCCAAATGTGAAAAATCTCATGTGGTCCACAGCTCACGTAGAAAAATGCAATGTCTGCA 3540
Qy      3429  |||||GlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeu 3448
Db      3541  |||||GGGTATCATTTATATATGAGACATGATCATCTACTCATGTTACAGTGATATCATGTTG 3600
Qy      3449  |||||GluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProPheCys 3468
Db      3601  |||||GAGGGTTCTCTGAGGAGTGTGTTGTTAGAAAAATGGAACATGGACATCACCTCTATTTCG 3660
Qy      3469  |||||ArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCys 3488
Db      3661  |||||AGAGCTGTCTGTGATTTTCATGTGCAAAATGGGGCATCTGCCAACCCCAATGCTGTGT 3720
Qy      3489  |||||SerCysProGluGlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeuProCys 3508

```

```

Db      3721  |||||TCCTGTCCAGAGGGCTGATGGGGCGCTCTGTGAGAGCAACCAATCTGCATTCTTCCCTGT 3780
Qy      3509  |||||LeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProGlyTyrThrGly 3528
Db      3781  |||||CTGAACGAGAGTGTGTGTGGCCCTTACCAGTGTGACTGCCCGCTGGCTGGACGGGG 3840
Qy      3529  |||||SerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArg 3548
Db      3841  |||||TCTCGCTGTATACACAGCTGTGTGCACTCTCCCTCTTAATGTGGAAATGTGTAGA 3900
Qy      3549  |||||ProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArg 3568
Db      3901  |||||CCAAACGATGTCACTGTCTTCTTCTTGGAGCGGACATACTGTTCGGGAAAGGAGG 3960
Qy      3569  |||||ThrGlyPhe 3571
Db      3961  |||||ACTGGGTTT 3969

RESULT 16
AX367105 3991 bp DNA linear PAT 16-FEB-2002
DEFINITION Sequence 24 from Patent WO0198354.
ACCESSION AX367105
VERSION AX367105.1 GI:18855307
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Griffin,J.A., Kallick,D.A., Tribouley,C.M., Yue,H., Nguyen,D.B.,
Tang,Y.T., Lal,P., Policky,J.L., Azimzai,Y., Lu,D.A., Graul,R.,
Yao,M.G., Burford,N., Hafalia,A.J., Baughn,M.R., Bandman,O.,
Patterson,C., Yang,J., Xu,Y., Warren,B.A., Ding,L. and
Sanjanwalla,M.S.
RECEPTORS
TITLE Receptors
JOURNAL Patent: WO 0198354-A 24 27-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
1. 3991
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3998749CB1"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3991
Score: 7471.00 Matches: 1327
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 3
Query Match: 37.41% Indels: 2
DB: 6 Gaps: 0

US-09-977-053-4 (1-3571) x AX367105 (1-3991)
Qy      2205  |||||GlyClnProProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGlu 2224
Db      1  |||||GGTGAACCACTAAGGTTGAGAAATGGCTTTCTGGAGCATACAACCTGGCAGGATCTTTGAG 60
Qy      2225  |||||SerGluValArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPheVal 2244
Db      61  |||||AGTGAAGTAGGATATCAGTGTAAACCCGGGCTATAGTCACTGATGTGTCTCTCGACTG 120
Qy      2245  |||||CysGln-AlaAsnArgHisTrpHisSerGluSerProLeuMetCysValProLeuAspCy 2264
Db      121  |||||TGCAAGGCCAATCGCCACTGGCACAGTGAATCCCTCTGATGTGTCTCTCGACTG 180
Qy      2264  |||||sGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPheGluValGlySe 2284
Db      181  |||||TGGAAACCTCCCGCATCCAGATGCTTCATGAAAGGAGAAACCTTTGAAGTAGGGTC 240
Qy      2284  |||||rLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTrpThrCy 2304

```

Db	241		CAAGGTT	CAGT	TTTTT	CTGT	TAAT	GAGGGT	TAT	GAGCT	TTG	TGGT	GACAGT	TTCT	TGG	CATG	300	
Qy	2304	sGlnLysSer	G	L	y	s	r	G	L	y	s	r	T	r	p	A	s	2324
Db	301	TCGAAAT	CT	TGG	CAAT	CGGAAT	AGAAGT	C	AAAT	CCAAAGT	GC	ATG	CT	G	CAAGT	CGCC	360	
Qy	2324	oGluPro	Pro	Leu	Leu	Glu	Asn	Gln	Leu	Val	Leu	Lys	Glu	Leu	Thr	Glu	2344	
Db	361	AGAGCG	CGCC	CT	CT	TGG	AAAC	CAG	CT	AGT	ATT	AAAG	GAGT	TTG	AC	CAC	420	
Qy	2344	lValThr	Ph	Ser	Cys	Al	ys	Glu	Gly	His	Val	Leu	Gln	Gly	Pro	Ser	2364	
Db	421	TGTGACA	TTT	CT	TGT	TAAGA	AGG	CA	TG	CT	CT	GC	AAAG	CGCC	CT	CT	480	
Qy	2364	uProSer	Gln	Gln	Trp	Asn	Asp	Ser	Pro	Val	Cys	Lys	Leu	Val	Leu	Cys	2384	
Db	481	GCCATC	CC	CA	CAAT	TGG	AAAT	GACT	CT	TT	CC	TG	TTG	TAA	GAT	GT	540	
Qy	2384	oProLeu	Leu	Ser	Ph	e	Gly	Val	Pro	Leu	Pro	Ser	Ser	Ala	Leu	His	2404	
Db	541	TC	CC	CT	AA	TT	TC	TT	CG	TT	CG	CT	CT	CT	CT	CT	600	
Qy	2404	lLysTrp	Ser	Cys	Val	Gly	Gly	Ph	e	Leu	Arg	Gly	Asn	Ser	Thr	Thr	2424	
Db	601	CAAGTAT	TC	TT	CT	TAG	CTGG	TTTT	TC	CT	TAAG	AGG	AAAT	TT	CT	AC	660	
Qy	2424	oAspGly	Thr	Trp	Ser	Ser	Pro	Leu	Pro	Glu	Cys	Val	Pro	Val	Glu	Cys	2444	
Db	661	TGATGG	CA	CT	CT	CT	CA	CT	GC	CA	GAAT	GT	TT	CC	AGT	AGAAT	720	
Qy	2444	uGluLeu	Pro	Asn	Gly	Leu	Leu	Asp	Val	Gln	Gly	Leu	Ala	Arg	Lys	Ser	2464	
Db	721	GGAAAT	TC	CC	CAAT	TGA	TG	TC	GAAG	GC	TT	GC	CT	TA	CT	CAG	780	
Qy	2464	rThrCys	Lys	Pro	Gly	Ph	e	Glu	Leu	Val	Gly	Asn	Thr	Thr	Thr	Lys	2484	
Db	781	TAC	CT	CG	CA	CC	AG	CT	TT	CAAT	TTGG	TGG	GAAT	TACT	ACC	AC	840	
Qy	2484	yHisTrp	Leu	Gly	Cys	Lys	Pro	Thr	Cys	Lys	Ala	Leu	Glu	Cys	Leu	Lys	2504	
Db	841	TC	ACT	GC	CT	TGG	AGG	AAAC	CA	CA	TG	TAAG	CC	CT	TAAG	CC	900	
Qy	2504	eLeuAsn	Gly	Lys	Ph	e	Ser	Trp	Thr	Asp	Leu	His	Trp	Gly	Gln	Thr	2524	
Db	901	TTTGAAT	TGG	CAAT	CT	CT	TAC	ACG	AC	CT	TAC	ACT	TG	GAC	AG	AC	960	
Qy	2524	sAsnArg	Lys	Ph	e	Arg	Leu	Glu	Gly	Pro	Ser	Ala	Leu	Thr	Cys	Leu	2544	
Db	961	CAAC	CG	AG	GG	CT	TT	CG	GT	CG	AA	GT	CC	AG	TC	CT	1020	
Qy	2544	pAspVal	Asp	Ala	Pro	Ser	Cys	Asn	Ala	Leu	His	Cys	Asp	Ser	Pro	Gln	2564	
Db	1021	GGAT	GT	AGA	TG	AC	CT	CT	TC	GAAT	GC	CA	CT	GT	GAAT	T	1080	
Qy	2564	nGly	Ph	e	Val	Glu	Gly	Ala	Asp	Trp	Ser	Trp	Gly	Ala	Leu	Leu	2584	
Db	1081	TGG	TTTT	TG	TAGA	AGG	TG	CAG	ATT	TAC	AGT	TAT	G	TG	CC	ATA	1140	
Qy	2584	oGly	Ph	e	Gln	Val	Ala	Gly	His	Ala	Met	Gln	Thr	Cys	Glu	Glu	2604	
Db	1141	TGGG	TTT	CA	GT	TGG	CT	GT	CA	CT	GC	AG	CC	CT	GT	GA	1200	
Qy	2604	rLysPro	Thr	Cys	Met	Pro	Leu	Asp	Cys	Gly	Leu	Pro	Pro	His	Leu	Asp	2624	
Db	1201	CAT	CC	CA	CA	AT	GT	AT	GC	CA	AT	AG	CT	GT	GC	CT	1260	
Qy	2624	sThrLys	Leu	Lys	Asp	Asp	Gln	Gly	Trp	Ph	e	Glu	Gln	Glu	Asp	Met	2644	
Db	1261	TACT	AACT	CA	AG	TG	AC	CG	GAAT	TTT	TG	AG	CA	AG	AG	CA	1320	
Qy	2644	oTrpVal	Thr	Pro	His	Pro	Trp	His	Leu	Gly	Ala	Val	Ala	Lys	Thr	Trp	2664	

1321	ATACGTGACTCTCTCAACCTCTCTTATCATTTGGGAGCAGTGGCTTAAACCTGGGAAATATC	1380
2664	rYsGiuSerProAlaThrHisSerSerAsnPheLeuTyrgLyThrMetValSerTyrrTh	2684
1381	AAAGGAGTCTCTGCTACACATTCATCAAACTTTTCGTATGGTATCCATGCTTTTCATACAC	1440
2684	rCysAsnProGlyTyrgLuleuLeuGlyAsnProValLeuIleCysGlnGluAspGlyTh	2704
1441	CTGTATATCCAGGATATGAACCTTCCTGGGGAAACCTCTGTGCTGATCTGCCAGGAAGATGGAAC	1500
2704	rTPAsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGly	2724
1501	TTGGATGGCAGTGCACCATCTCGCATTTCAATTGAATGTGACTTTCCTACTCTCTCTGA	1560
2724	uAsnGlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrrSerCysLy	2744
1561	AAATGGCTTTTTCGTTTTACAGAGACTAGCATGGGAGTGTCTGTGCAGTATAGCTGTAA	1620
2744	sProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgIysTrpSe	2764
1621	ACCTGGACACATTTCTAGCAGGCTCTGACTTAAGCTTTTGTCTAGAGAAATAGAAAGTGGAG	1680
2764	rGlyAlaSerProArgCysGluAlaIleSerCysIysIysProAsnProValMetAsnGly	2784
1681	TGTGTGCTCCCCAGCTGTGAAGCCATTTTCATGCMAAAAGCCAAATTCAGTCAATGAATGG	1740
2784	ySerIleYsGlySerAsnTyrrThrTyrrLeuSerThrLeuTyrrTyrgLyuCysAspProGly	2804
1741	ATCANTCMAAGGAGACACTACACATCTCTGACACAGTGTACTATGAGTGTGACCCCGG	1800
2804	yTyrrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspIysAsnTrpAspGluAs	2824
1801	ATATGTCTGTAATGGCACTGAGAGGAGAACATGCCAGGATGACAAAAAATCGGGATGAGGA	1860
2824	pGluProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyGlnVa	2844
1861	TGAGCCCATTTGCATTTCTGTGGACTGCACTGTCACTCCACAGTCTCAGCCCAATGGCCAGGT	1920
2844	lArgGlyAspGluTyrrPheGlnIysGluIleGluTyrrThrCysAsnGluGlyPheLe	2864
1921	GAGAGGAGACGATACACATTCMAAAGAGATTGAATACACTTGCATGAAGGGTTCCTT	1980
2864	uLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrPr	2884
1981	GCCTGAGGAGCCAGGAGTCGGGTTCCTTGCCATGGAAATGGAGTTGGAGTGGAGCCACTCC	2040
2884	oAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGly	2904
2041	CGACTGTGTGCTGTGAGATGTGCCACCCCGGCACAACTGGCCCAATGGGGTGACGGAGG	2100
2904	yLeuAspTyrrGlyPheMetIysGluValThrPheHisCysHisGluGlyTyrrIleLeuHi	2924
2101	CCCTGGACTATGGCTTCATGAGGAGTAACATTCCTCACTGTCAAGGGCTACATCTTGCA	2160
2924	sGlyAlaProLyLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCy	2944
2161	CGGTGCTCCAAAACCTCACCTGTCACTGATGCAACTGGGATGCAAGATTCCTCTCTGT	2220
2944	sIysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSe	2964
2221	TAAACCACTCACTGTGGACCTCTGAGATCTTGCCCAATGGTTTTCCCTAATGGTTTTTC	2280
2964	rPheIleHisGlyIleGlyHisIleGlnTyrrGlnCysPheProGlyTyrrLyLeuHisGlyAs	2984
2281	CTTTATTCATGGGGCCATATACAGTATCAGTCTTCCTGTTTATTAAGCTCCATCGAATA	2340
2984	nSerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuPr	3004
2341	TTCATCAGAAGGTGCTCTCTCAAATGGCTCCTGGAGTGGCAGCTCACTTCTCTGCTGCC	2400
3004	oCysArgCysSerThrProValIleGluTyrrGlyThrValAsnGlyThrAspPheAspCy	3024
2401	TTGCAGATGTTCACACCACTGAATGAATATGAACCTGTCAATGGGACAGATTTGACTG	2460

```
3024 sGlyLysAlaAaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluI 3044
3045 |
3046 |
3047 |
3048 |
3049 |
3050 |
3051 |
3052 |
3053 |
3054 |
3055 |
3056 |
3057 |
3058 |
3059 |
3060 |
3061 |
3062 |
3063 |
3064 |
3065 |
3066 |
3067 |
3068 |
3069 |
3070 |
3071 |
3072 |
3073 |
3074 |
3075 |
3076 |
3077 |
3078 |
3079 |
3080 |
3081 |
3082 |
3083 |
3084 |
3085 |
3086 |
3087 |
3088 |
3089 |
3090 |
3091 |
3092 |
3093 |
3094 |
3095 |
3096 |
3097 |
3098 |
3099 |
3100 |
3101 |
3102 |
3103 |
3104 |
3105 |
3106 |
3107 |
3108 |
3109 |
3110 |
3111 |
3112 |
3113 |
3114 |
3115 |
3116 |
3117 |
3118 |
3119 |
3120 |
3121 |
3122 |
3123 |
3124 |
3125 |
3126 |
3127 |
3128 |
3129 |
3130 |
3131 |
3132 |
3133 |
3134 |
3135 |
3136 |
3137 |
3138 |
3139 |
3140 |
3141 |
3142 |
3143 |
3144 |
3145 |
3146 |
3147 |
3148 |
3149 |
3150 |
3151 |
3152 |
3153 |
3154 |
3155 |
3156 |
3157 |
3158 |
3159 |
3160 |
3161 |
3162 |
3163 |
3164 |
3165 |
3166 |
3167 |
3168 |
3169 |
3170 |
3171 |
3172 |
3173 |
3174 |
3175 |
3176 |
3177 |
3178 |
3179 |
3180 |
3181 |
3182 |
3183 |
3184 |
3185 |
3186 |
3187 |
3188 |
3189 |
3190 |
3191 |
3192 |
3193 |
3194 |
3195 |
3196 |
3197 |
3198 |
3199 |
3200 |
3201 |
3202 |
3203 |
3204 |
3205 |
3206 |
3207 |
3208 |
3209 |
3210 |
3211 |
3212 |
3213 |
3214 |
3215 |
3216 |
3217 |
3218 |
3219 |
3220 |
3221 |
3222 |
3223 |
3224 |
3225 |
3226 |
3227 |
3228 |
3229 |
3230 |
3231 |
3232 |
3233 |
3234 |
3235 |
3236 |
3237 |
3238 |
3239 |
3240 |
3241 |
3242 |
3243 |
3244 |
3245 |
3246 |
3247 |
3248 |
3249 |
3250 |
3251 |
3252 |
3253 |
3254 |
3255 |
3256 |
3257 |
3258 |
3259 |
3260 |
3261 |
3262 |
3263 |
3264 |
3265 |
3266 |
3267 |
3268 |
3269 |
3270 |
3271 |
3272 |
3273 |
3274 |
3275 |
3276 |
3277 |
3278 |
3279 |
3280 |
3281 |
3282 |
3283 |
3284 |
3285 |
3286 |
3287 |
3288 |
3289 |
3290 |
3291 |
3292 |
3293 |
3294 |
3295 |
3296 |
3297 |
3298 |
3299 |
3300 |
3301 |
3302 |
3303 |
3304 |
3305 |
3306 |
3307 |
3308 |
3309 |
3310 |
3311 |
3312 |
3313 |
3314 |
3315 |
3316 |
3317 |
3318 |
3319 |
3320 |
3321 |
3322 |
3323 |
3324 |
3325 |
3326 |
3327 |
3328 |
3329 |
3330 |
3331 |
3332 |
3333 |
3334 |
3335 |
3336 |
3337 |
3338 |
3339 |
3340 |
3341 |
3342 |
3343 |
3344 |
3345 |
3346 |
3347 |
3348 |
3349 |
3350 |
3351 |
3352 |
3353 |
3354 |
3355 |
3356 |
3357 |
3358 |
3359 |
3360 |
3361 |
3362 |
3363 |
3364 |
3365 |
3366 |
3367 |
3368 |
3369 |
3370 |
3371 |
3372 |
3373 |
3374 |
3375 |
3376 |
3377 |
3378 |
3379 |
3380 |
3381 |
3382 |
3383 |
3384 |
3385 |
3386 |
3387 |
3388 |
3389 |
3390 |
3391 |
3392 |
3393 |
3394 |
3395 |
3396 |
3397 |
3398 |
3399 |
3400 |
3401 |
3402 |
3403 |
3404 |
3405 |
3406 |
3407 |
3408 |
3409 |
3410 |
3411 |
3412 |
3413 |
3414 |
3415 |
3416 |
3417 |
3418 |
3419 |
3420 |
3421 |
3422 |
3423 |
3424 |
3425 |
3426 |
3427 |
3428 |
3429 |
3430 |
3431 |
3432 |
3433 |
3434 |
3435 |
3436 |
3437 |
3438 |
3439 |
3440 |
3441 |
3442 |
3443 |
3444 |
3445 |
3446 |
3447 |
3448 |
3449 |
3450 |
3451 |
3452 |
3453 |
3454 |
3455 |
3456 |
3457 |
3458 |
3459 |
3460 |
3461 |
3462 |
3463 |
3464 |
3465 |
3466 |
3467 |
3468 |
3469 |
3470 |
3471 |
3472 |
3473 |
3474 |
3475 |
3476 |
3477 |
3478 |
3479 |
3480 |
3481 |
3482 |
3483 |
3484 |
3485 |
3486 |
3487 |
3488 |
3489 |
3490 |
3491 |
3492 |
3493 |
3494 |
3495 |
3496 |
3497 |
3498 |
3499 |
3500 |
3501 |
3502 |
3503 |
3504 |
3505 |
3506 |
3507 |
3508 |
3509 |
3510 |
3511 |
3512 |
3513 |
3514 |
3515 |
3516 |
3517 |
3518 |
3519 |
3520 |
3521 |
3522 |
3523 |
3524 |
3525 |
3526 |
3527 |
3528 |
3529 |
3530 |
3531 |
3532 |
3533 |
3534 |
3535 |
3536 |
3537 |
3538 |
3539 |
3540 |
3541 |
3542 |
3543 |
3544 |
3545 |
3546 |
3547 |
3548 |
3549 |
3550 |
3551 |
3552 |
3553 |
3554 |
3555 |
3556 |
3557 |
3558 |
3559 |
3560 |
3561 |
3562 |
3563 |
3564 |
3565 |
3566 |
3567 |
3568 |
3569 |
3570 |
3571 |
3572 |
3573 |
3574 |
3575 |
3576 |
3577 |
3578 |
3579 |
3580 |
3581 |
3582 |
3583 |
3584 |
3585 |
3586 |
3587 |
3588 |
3589 |
3590 |
3591 |
3592 |
3593 |
3594 |
3595 |
3596 |
3597 |
3598 |
3599 |
3600 |
3601 |
3602 |
3603 |
3604 |
3605 |
3606 |
3607 |
3608 |
3609 |
3610 |
3611 |
3612 |
3613 |
3614 |
3615 |
3616 |
3617 |
3618 |
3619 |
3620 |
3621 |
3622 |
3623 |
3624 |
3625 |
3626 |
3627 |
3628 |
3629 |
3630 |
3631 |
3632 |
3633 |
3634 |
3635 |
3636 |
3637 |
3638 |
3639 |
3640 |
3641 |
3642 |
3643 |
3644 |
3645 |
3646 |
3647 |
3648 |
3649 |
3650 |
3651 |
3652 |
3653 |
3654 |
3655 |
3656 |
3657 |
3658 |
3659 |
3660 |
3661 |
3662 |
3663 |
3664 |
3665 |
3666 |
3667 |
3668 |
3669 |
3670 |
3671 |
3672 |
3673 |
3674 |
3675 |
3676 |
3677 |
3678 |
3679 |
3680 |
3681 |
3682 |
3683 |
3684 |
3685 |
3686 |
3687 |
3688 |
3689 |
3690 |
3691 |
3692 |
3693 |
3694 |
3695 |
3696 |
3697 |
3698 |
3699 |
3700 |
3701 |
3702 |
3703 |
3704 |
3705 |
3706 |
3707 |
3708 |
3709 |
3710 |
3711 |
3712 |
3713 |
3714 |
3715 |
3716 |
3717 |
3718 |
3719 |
3720 |
3721 |
3722 |
3723 |
3724 |
3725 |
3726 |
3727 |
3728 |
3729 |
3730 |
3731 |
3732 |
3733 |
3734 |
3735 |
3736 |
3737 |
3738 |
3739 |
3740 |
3741 |
3742 |
3743 |
3744 |
3745 |
3746 |
3747 |
3748 |
3749 |
3750 |
3751 |
3752 |
3753 |
3754 |
3755 |
3756 |
3757 |
3758 |
3759 |
3760 |
3761 |
3762 |
3763 |
3764 |
3765 |
3766 |
3767 |
3768 |
3769 |
3770 |
3771 |
3772 |
3773 |
3774 |
3775 |
3776 |
3777 |
3778 |
3779 |
3780 |
3781 |
3782 |
3783 |
3784 |
3785 |
3786 |
3787 |
3788 |
3789 |
3790 |
3791 |
3792 |
3793 |
3794 |
3795 |
3796 |
3797 |
3798 |
3799 |
3800 |
3801 |
3802 |
3803 |
3804 |
3805 |
3806 |
3807 |
3808 |
3809 |
3810 |
3811 |
3812 |
3813 |
3814 |
3815 |
3816 |
3817 |
3818 |
3819 |
3820 |
3821 |
3822 |
3823 |
3824 |
3825 |
3826 |
3827 |
3828 |
3829 |
3830 |
3831 |
3832 |
3833 |
3834 |
3835 |
3836 |
3837 |
3838 |
3839 |
3840 |
3841 |
3842 |
3843 |
3844 |
3845 |
3846 |
3847 |
3848 |
3849 |
3850 |
3851 |
3852 |
3853 |
3854 |
3855 |
3856 |
3857 |
3858 |
3859 |
3860 |
3861 |
3862 |
3863 |
3864 |
3865 |
3866 |
3867 |
3868 |
3869 |
3870 |
3871 |
3872 |
3873 |
3874 |
3875 |
3876 |
3877 |
3878 |
3879 |
3880 |
3881 |
3882 |
3883 |
3884 |
3885 |
3886 |
3887 |
3888 |
3889 |
3890 |
3891 |
3892 |
3893 |
3894 |
3895 |
3896 |
3897 |
3898 |
3899 |
3900 |
3901 |
3902 |
3903 |
3904 |
3905 |
3906 |
3907 |
3908 |
3909 |
3910 |
3911 |
3912 |
3913 |
3914 |
3915 |
3916 |
3917 |
3918 |
3919 |
3920 |
3921 |
3922 |
3923 |
3924 |
3925 |
3926 |
3927 |
3928 |
3929 |
3930 |
3931 |
3932 |
3933 |
3934 |
3935 |
3936 |
3937 |
3938 |
3939 |
3940 |
3941 |
3942 |
3943 |
3944 |
3945 |
3946 |
3947 |
3948 |
3949 |
3950 |
3951 |
3952 |
3953 |
3954 |
3955 |
3956 |
3957 |
3958 |
3959 |
3960 |
3961 |
3962 |
3963 |
3964 |
3965 |
3966 |
3967 |
3968 |
3969 |
3970 |
3971 |
3972 |
3973 |
3974 |
3975 |
3976 |
3977 |
3978 |
3979 |
3980 |
3981 |
3982 |
3983 |
3984 |
3985 |
3986 |
3987 |
3988 |
3989 |
3990 |
3991 |
3992 |
3993 |
3994 |
3995 |
3996 |
3997 |
3998 |
3999 |
4000 |
```

```
3384 sArgGluGlyPheLeuLeuGlnGlyHisGlyIleThrCysAsnProAspGluThrTr 3404
3385 |
3386 |
3387 |
3388 |
3389 |
3390 |
3391 |
3392 |
3393 |
3394 |
3395 |
3396 |
3397 |
3398 |
3399 |
3400 |
3401 |
3402 |
3403 |
3404 |
3405 |
3406 |
3407 |
3408 |
3409 |
3410 |
3411 |
3412 |
3413 |
3414 |
3415 |
3416 |
3417 |
3418 |
3419 |
3420 |
3421 |
3422 |
3423 |
3424 |
3425 |
3426 |
3427 |
3428 |
3429 |
3430 |
3431 |
3432 |
3433 |
3434 |
3435 |
3436 |
3437 |
3438 |
3439 |
3440 |
3441 |
3442 |
3443 |
3444 |
3445 |
3446 |
3447 |
3448 |
3449 |
3450 |
3451 |
3452 |
3453 |
3454 |
3455 |
3456 |
3457 |
3458 |
3459 |
3460 |
3461 |
3462 |
3463 |
3464 |
3465 |
3466 |
3467 |
3468 |
3469 |
3470 |
3471 |
3472 |
3473 |
3474 |
3475 |
3476 |
3477 |
3478 |
3479 |
3480 |
3481 |
3482 |
3483 |
3484 |
3485 |
3486 |
3487 |
3488 |
3489 |
3490 |
3491 |
3492 |
3493 |
3494 |
3495 |
3496 |
3497 |
3498 |
3499 |
3500 |
3501 |
3502 |
3503 |
3504 |
3505 |
3506 |
3507 |
3508 |
3509 |
3510 |
3511 |
3512 |
3513 |
3514 |
3515 |
3516 |
3517 |
3518 |
3519 |
3520 |
3521 |
3522 |
3523 |
3524 |
3525 |
3526 |
3527 |
3528 |
3529 |
3530 |
3531 |
3532 |
3533 |
3534 |
3535 |
3536 |
3537 |
3538 |
3539 |
3540 |
3541 |
3542 |
3543 |
3544 |
3545 |
3546 |
3547 |
3548 |
3549 |
3550 |
3551 |
3552 |
3553 |
3554 |
3555 |
3556 |
3557 |
3558 |
3559 |
3560 |
3561 |
3562 |
3563 |
3564 |
3565 |
3566 |
3567 |
3568 |
3569 |
3570 |
3571 |
3572 |
3573 |
3574 |
3575 |
3576 |
3577 |
3578 |
3579 |
3580 |
3581 |
3582 |
3583 |
3584 |
3585 |
3586 |
3587 |
3588 |
3589 |
3590 |
3591 |
3592 |
3593 |
3594 |
3595 |
3596 |
3597 |
3598 |
3599 |
3600 |
3601 |
3602 |
3603 |
3604 |
3605 |
3606 |
3607 |
3608 |
3609 |
3610 |
3611 |
3612 |
3613 |
3614 |
3615 |
3616 |
3617 |
3618 |
3619 |
3620 |
3621 |
3622 |
3623 |
3624 |
3625 |
3626 |
3627 |
3628 |
3629 |
3630 |
3631 |
3632 |
3633 |
3634 |
3635 |
3636 |
3637 |
3638 |
3639 |
3640 |
3641 |
3642 |
3643 |
3644 |
3645 |
3646 |
3647 |
3648 |
3649 |
3650 |
3651 |
3652 |
3653 |
3654 |
3655 |
3656 |
3657 |
3658 |
3659 |
3660 |
3661 |
3662 |
3663 |
3664 |
3665 |
3666 |
3667 |
3668 |
3669 |
3670 |
3671 |
3672 |
3673 |
3674 |
3675 |
3676 |
3677 |
3678 |
3679 |
3680 |
3681 |
3682 |
3683 |
3684 |
3685 |
3686 |
3687 |
3688 |
3689 |
3690 |
3691 |
3692 |
3693 |
3694 |
3695 |
3696 |
3697 |
3698 |
3699 |
3700 |
3701 |
3702 |
3703 |
3704 |
3705 |
3706 |
3707 |
3708 |
3709 |
3710 |
3711 |
3712 |
3713 |
3714 |
3715 |
3716 |
3717 |
3718 |
3719 |
3720 |
3721 |
3722 |
3723 |
3724 |
3725 |
3726 |
3727 |
3728 |
3729 |
3730 |
3731 |
3732 |
3733 |
3734 |
3735 |
3736 |
3737 |
3738 |
3739 |
3740 |
3741 |
3742 |
3743 |
3744 |
3745 |
3746 |
3747 |
3748 |
3749 |
3750 |
3751 |
3752 |
3753 |
3754 |
3755 |
3756 |
3757 |
3758 |
3759 |
3760 |
3761 |
3762 |
3763 |
3764 |
3765 |
3766 |
3767 |
3768 |
3769 |
3770 |
3771 |
3772 |
3773 |
3774 |
3775 |
3776 |
3777 |
3778 |
3779 |
3780 |
3781 |
3782 |
3783 |
3784 |
3785 |
3786 |
3787 |
3788 |
3789 |
3790 |
3791 |
3792 |
3793 |
3794 |
3795 |
3796 |
3797 |
3798 |
3799 |
3800 |
3801 |
3802 |
3803 |
3804 |
3805 |
3806 |
3807 |
3808 |
3809 |
3810 |
3811 |
3812 |
3813 |
3814 |
3815 |
3816 |
3817 |
3818 |
3819 |
3820 |
3821 |
3822 |
3823 |
3824 |
3825 |
3826 |
3827 |
3828 |
3829 |
3830 |
3831 |
3832 |
3833 |
3834 |
3835 |
3836 |
3837 |
3838 |
3839 |
3840 |
3841 |
3842 |
3843 |
3844 |
3845 |
3846 |
3847 |
3848 |
3849 |
3850 |
3851 |
3852 |
3853 |
3854 |
3855 |
3856 |
3857 |
3858 |
3859 |
3860 |
3861 |
3862 |
3863 |
3864 |
3865 |
3866 |
3867 |
3868 |
3869 |
3870 |
3871 |
3872 |
3873 |
3874 |
3875 |
3876 |
3877 |
3878 |
3879 |
3880 |
3881 |
3882 |
3883 |
3884 |
3885 |
3886 |
3887 |
3888 |
3889 |
3890 |
3891 |
3892 |
3893 |
3894 |
3895 |
3896 |
3897 |
3898 |
3899 |
3900 |
3901 |
3902 |
3903 |
3904 |
3905 |
3906 |
3907 |
3908 |
3909 |
3910 |
3911 |
3912 |
3913 |
3914 |
3915 |
3916 |
3917 |
3918 |
3919 |
3920 |
3921 |
3922 |
3923 |
3924 |
3925 |
3926 |
3927 |
3928 |
3929 |
3930 |
3931 |
3932 |
3933 |
3934 |
3935 |
3936 |
3937 |
3938 |
3939 |
3940 |
3941 |
3942 |
3943 |
3944 |
3945 |
3946 |
3947 |
3948 |
3949 |
3950 |
3951 |
3952 |
3953 |
3954 |
3955 |
3956 |
3957 |
3958 |
3959 |
3960 |
3961 |
3962 |
3963 |
3964 |
3965 |
3966 |
3967 |
3968 |
3969 |
3970 |
3971 |
3972 |
3973 |
3974 |
3975 |
3976 |
3977 |
3978 |
3979 |
3980 |
3981 |
3982 |
3983 |
3984 |
3985 |
3986 |
3987 |
3988 |
3989 |
3990 |
3991 |
3992 |
3993 |
3994 |
3995 |
3996 |
3997 |
3998 |
3999 |
4000 |
```

RESULT 17

```
BD127948
LOCUS BD127948 4088 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127948
VERSION BD127948.1 GI:23222893
KEYWORDS JP 2002017375-A/3379.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4088)
AUTHORS Rukaryoka; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota, T., Nishikawa, T., Isogai, T., Hayashi, K., Ishii, S., Kawai, Y.,
Wakamatsu, A., Sugiyama, T., Nagai, K., Kojima, S., Otsuki, T. and
Koga, H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL HELIX RESEARCH INSTITUTE
COMMENT OS Homo sapiens (human)
PW JP 2002017375-A/3379
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA, TETSUO NISHIKAWA, TAKAO ISOGAI, KOJI HAYASHI, SHIZUKO
PI ISHII,
PI YURI KAWAI, AI WAKAMATSU, TOMOYASU SUGIYAMA, KEIICHI NAGAI, PI
SHINTCHI KOJIMA,
PI TETSUJI OTSUKI, HISASHI KOGA
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
1..(1545)
1..4088
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
```

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 4088
 Score: 6950.00 Matches: 1232
 Percent Similarity: 99.20% Conservative: 1
 Best Local Similarity: 99.12% Mismatches: 10
 Query Match: 34.80% Indels: 1
 DB: 6 Gaps: 0

US-09-977-053-4 (1-3571) x BD127948 (1-4088)

Qy 2329 LeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSer 2348
 Db 1 ATGGAAACACGCTAGTATTAAAGAGTGTGACCCAGGAGTAGGAGTGTGACATTTTC 60
 Qy 2349 CysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGln 2368
 Db 61 TGTAAAGAAAGGCGCATGTCTGCAAGGCCCTCTGTCTGTAAGTGTCTGCAATGCTTGCATCCAGCA 120
 Qy 2369 TrpAspSerPheProValCysLysIleValLeuCysThrThrProProLeuIleSer 2388
 Db 121 TGGAAATGACTCTTTCCCTGTTTGAAGATGTCTTGTGATCCACCTCCCTTAATTTCC 180
 Qy 2389 PheGlyValProIleProSerSerAlaLeuHisPheGlySerThrValLysTyrSerCys 2408
 Db 181 TTTGGTGTCCCATCTCTCTCTGCTCTTCATTTTGAAGTACTGTCAAGTATCTGT 240
 Qy 2409 ValGlyGlyPheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrp 2428
 Db 241 GTAGTGGGTTTTCTTAAGAGGAAATCTACCAACCTCTGCCAACCTGTATGGCACCTGG 300
 Qy 2429 SerSerProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsn 2448
 Db 301 AGCTCTCCACTGCCAGAAATGTCTCAGTAGAATGTCCCCAACCTGAGGAAATCCCCAAT 360
 Qy 2449 GlyIleIleAspValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysLysPro 2468
 Db 361 GGAATCATGTATGTGCAAGGCTTGGCTATCTCAGCAGACTCTCTATACCTGCAAGCCA 420
 Qy 2469 GlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGly 2488
 Db 421 GGCCTTGAAATGTGGGGAATACTACCACCTTTGTGGGAAATGGTCACTGGCTTGG 480
 Qy 2489 GlyLysProThrCysLysAlaIleGluCysLysLysProLysGluIleLeuAsnGlyLys 2508
 Db 481 GGAATAACCAATGATAAGCCATGTAGTGCTTGAAACCCAAAGGAGATTTTGAATGGCAAA 540
 Qy 2509 PheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPhe 2528
 Db 541 TTCTTTACACGGACCTACACTATGGACAGACCGTTACTTACTCTTGCAACGGAGCTTT 600
 Qy 2529 ArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAla 2548
 Db 601 CGGCTCGAAGGTCCCAAGTGCCTTGACCTGTTTATAGACAGAGTGTATGGGATGTAGATGCC 660
 Qy 2549 ProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGlu 2568
 Db 661 CCATCTTGAATGCCATCCACTGTGATTTCCCAACCCATGTAAGATGGTGTGTAGNA 720
 Qy 2569 GlyAlaAspTyrSerTyrGlyAlaIleIleTyrSerCysPhePheGlyPheGlnVal 2588
 Db 721 GGTCCAGATTACAGCTATGGTGGCCATATCATCATCTACAGTGTCTTCCCTGGGTTTCAGGTG 780
 Qy 2589 AlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerSerIleProThrCys 2608
 Db 781 GCTGGTCATGTCATGCAGACCTGTGAAGAGTCAGGATGGTCAAGTTCATCCCAACATGT 840
 Qy 2609 MetProIleAspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLys 2628
 Db 841 ATGCCAATAGACTGTGGCTTCCCTCTCATATAGATTTTGGAGACTGTACTTAACCTCAAA 900
 Qy 2629 AspAspGlnGlyTyrPheGluGlnGluAspAspMetMetGluValProTyrValThrPro 2648
 Db 901 GATGACACGGGATATTATTGACGACGAAGACCATGATGGAAATTTCCATATGTGACTCT 960

Qy 2649 HisProProTyrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerPro 2668
 Db 961 CACCTCTCTTATCATTTTGGAGCAGCGCTAAACCTGGGAAATACAAAGAGTCTCTCT 1020
 Qy 2669 AlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGly 2688
 Db 1021 GCTACACATTCTAACAACCTTCTGTATGTATGATACCATGTTTCTATACACCTGTATCCAGA 1080
 Qy 2689 TyrGluLeuLeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrpAsnGlySer 2708
 Db 1081 TATGAACCTTCTGGGNAACCTCTGTCTGATCTGCCAGGAAGATGGAACCTTGGAAATGGCAGT 1140
 Qy 2709 AlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeu 2728
 Db 1141 GCACCATCTCTGCATTTCAATTTGAATGTGACTTGTCTACTGTCTCTGAAATATGGCTTTTGG 1200
 Qy 2729 ArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCysLysProGlyHisIle 2748
 Db 1201 CGTTTACAGAGACTAGCATGTGGAAAGTCTGTGTGCAATATAGCTGTAAACCTGGACACATT 1260
 Qy 2749 LeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaSerPro 2768
 Db 1261 CTAGCAGGCTCTGACTTAAGGCTTTGTCTAGAGAATAGAAAGTGGAGTGTGCTCTCCCA 1320
 Qy 2769 ArgCysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGly 2788
 Db 1321 CGCTGTGAAGCCATTTCTATGCANAAAGCCAAATCCAGTCATGAATGGATCCATCAAGGA 1380
 Qy 2789 SerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAspProGlyTyrValLeuAsn 2808
 Db 1381 AGCAACTACATACCTAGGACAGCTGTGTACTATGTAGTGTGCCCCCGATATGTGTGAT 1440
 Qy 2809 GlyThrGluArgThrCysGlnAspAspLysAsnTrpAspGluAspGluProIleCys 2828
 Db 1441 GGCACTGAGAGAGAAACATGCCAGGATGACAA-NACTGGATGAGGATGAGGCCATTTGC 1499
 Qy 2829 IleProValAspCysSerSerProValSerAlaAsnGlyGlnValArgGlyAspGlu 2848
 Db 1500 ATTCCTGTGGATCGAGTTTCACTCCAGCTCTCAGCCAAATGCCAGGTGAGAGAGACGAG 1559
 Qy 2849 TyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLeuLeuGluGlyAla 2868
 Db 1560 TACATTTCCAAAGAGATTTGAATACACTTGCATGAAGGGTCTCTGCTGAGGGAGCC 1619
 Qy 2869 ArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAspCysValPro 2888
 Db 1620 AGGAGTGGGTGTGTTGTTGCAATGGAAATGGAGTGGAGGCCATCTCCGACCTGTGTGCT 1679
 Qy 2889 ValArgCysAlaThrProProGluLeuAlaAsnGlyValThrGluGlyLeuAspTyrGly 2908
 Db 1680 GTCAGATGTGCCACCCGCCCAACTGGCCAAATGGGGTGAACGAGGCTTGACATATGCG 1739
 Qy 2909 PheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHisGlyAlaProLys 2928
 Db 1740 TTCATGAAGAGTAACATTCATCTCATGAGGCTACATCTTGCACGGTCTCCCAAA 1799
 Qy 2929 LeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsn 2948
 Db 1800 CTCACCTCTCAGTCAGATGGCAACTGGGATGAGAGATTCCTCTCTGTAACACAGTCAAC 1859
 Qy 2949 CysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGly 2968
 Db 1860 TGTGACCTCTCTGAAGATCTTGGCCANGGTTTCCCTAATGGTTTTCTTTATTCATGGG 1919
 Qy 2969 GlyHisIleGlnTyrGlnCysPhePheProGlyTyrLysLeuHisGlyAsnSerSerArgArg 2988
 Db 1920 GGCATATACAGTATCAGTCTCTTCTGTTATAGCTCCATGGAAATTCATCAAGAGG 1979
 Qy 2989 CysLeuSerAsnGlySerTrpSerGlySerProSerCysLeuProCysArgCysSer 3008
 Db 1980 TGCCTCTCCAAATGGCTCTGGAGTGGCAGCTCACCTTCTCTGCTGCTGCTGCTGCTGCT 2039

3009 ThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaala 3028
 3040 ACACCGAGTAATGATATGAACTGTCTCAATGGGACAGATTTTGAATCTGGAAAGCGAGCC 2099
 3029 ArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAla 3048
 2100 CGGATTCAAGTCTTCAAAAGGCTTCAAGCTCTAGGACTTTCTGAATCACTGTGAAGCC 2159
 3049 AspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuPro 3068
 2160 GATGGCCAGTGGAGCTCTGGGTTCCCCACCTGTGAACACACTTCTTGGGTCTCTTCCA 2219
 3069 MetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGluAsnValIleThrTyr 3088
 2220 ATGATACCAATGGCTTCATCAGTGAACACGACTTCTTGGAAAGGAATGTGATNACTTAC 2279
 3089 SerCysArgSerGlyTyrValIleGlnLysSerAspLeuIleCysThrGluLysGly 3108
 2280 AGCTGCAGGCTCTGATATGTCATCAAGGAGGAGTTCAGATCTGAGTGTGTACAGAGAAAGG 2339
 3109 ValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerVal 3128
 2340 GTATGGAGCCAGCCCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCCCCACCGTCTGTC 2399
 3129 AlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCys 3148
 2400 GCCAATCCAGTGGCACTTGGAGAGGACACACCTTATGAAAGTGAAGTGAATCTCGATGT 2459
 3149 LeuGluGlyTyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyArg 3168
 2460 CTGGAGGTTATACGATGGATACAGATACAGATACCAATCACTGTCTGAGAAAGATGGTCGC 2519
 3169 TrpPheProGluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleThr 3188
 2520 TGGTTCCTCAGAGAAATCTCTCGAGTCTCTAAATAAATGTCTCTCCCGGAAATCAATACA 2579
 3189 HisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAla 3208
 2580 CATATACCTTGATCATGGGAGCATTCAGTGTGTGATAGGCAGTTTCTGTGTCAATGTCA 2639
 3209 GluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGlu 3228
 2640 GAAGGTTATATCTTTGAGTCAAGTTAAATATACATATCAGTATGTCAGCTTGTGAACCTGGGAG 2699
 3229 ProProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGlu 3248
 2700 CCACCAATCTCCGATGAATCTTCAGTCCAGTCTCTTGTGGAAACCTGAAAGTCCAGAA 2759
 3249 HisGlyPheValValGlySerLysThrPheGluSerThrIleIleTyrGlnCysGlu 3268
 2760 CATGGATTGTGGTGGCAGTAAATATACACTTTGAAAGCACATATTTATCATGTGTGAG 2819
 3269 ProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSer 3288
 2820 CCTGGCTATGAATAGAGGGGAAACAGGGAACGTGTCTGCCAGGAGAACACAGACAGTGGAGT 2879
 3289 GlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGly 3308
 2880 GGAAGGTTGGCAATATGCAAGAGACAGGTTGTGAAATCTCACTTGAATTTCTCAATGGG 2939
 3309 LysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGly 3328
 2940 AAGCTGACATTTGAACAGGACGACCTGGACCCCACTGGTATATTTCTTCAACAGAGGC 2999
 3329 TyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisPro 3348
 3000 TACAGTCTTGAAGGGCCATCTGAGGCACACTGACAGAAAATGGAACCTGGAGGCCACCCA 3059
 3349 ValProLeuCysLysProAsnProCysProValProPheValIleProGluAsnAlaLeu 3368
 3060 GTCCCTCTCTGCAAAACCAATTCATGCCCCCTTCTTTGTGATTCCTCCGAGAAATGCTCTG 3119
 3369 LeuSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPhe 3388

3120 CTGTCTCAAAAGGAGTTTTATGTTGATCAGAAATGTTCCATCAAAATGTAGGAAAGGTTTT 3179
 3389 LeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSer 3408
 3180 CTGTGAGGGCCACCGCATCATTTACCTGCAACCCCGACGAGCGTGGACACAGCAAGC 3239
 3409 AlalysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArg 3428
 3240 GCCAATGTGMAAANAATCTCATGTGTTCCACAGCTCAGCTAGAAAATGCAATTCCTCGA 3299
 3429 GlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeu 3448
 3300 GGCGTACATATTAATATGGAGACATGATCACTACTCATGTTACAGTGGATACATGTTG 3359
 3449 GluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCys 3468
 3360 GAGGGTTTCAGAGGAGTGTGTTTGAANAATGGAACATGGACATCACTCTCTATTTTGC 3419
 3469 ArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCys 3488
 3420 AGACTGTCTGTGATTTCCATGTCAAGATGGGGGCATCTGCCAACGCCCAATGCTTGT 3479
 3489 SerCysProGluGlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeuProCys 3508
 3480 TCCTGTCCAGAGGCTGAGTGGGGCGCTCTGTGAAGAACCAATCAGCATCTCTCCCTGT 3539
 3509 LeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGlyTyrThrGly 3528
 3540 CTGAACGGAGTCTGCTGTGTGGCCCCCTTACCAAGTGTGACTGCCCGCTGGCGAGGGG 3599
 3529 SerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArg 3548
 3600 TCTCGCTGTCTAGCAGCTGTGTTCAGCTTCCCTGCTTAATGTGGAAATGTGTGAAGA 3659
 3549 ProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArg 3568
 3660 CCAACCGATGTCACTGTCTTCTTCTTGGAGCGGACATAACTGTTCAGGAAAAAGGAGG 3719
 3569 ThrGlyPhe 3571
 3720 ACTGGGTTTT 3728

RESULT 18

AK075234 4088 bp mRNA linear PRI 03-SEP-2002
 LOCUS Homo sapiens cDNA FLJ90753 fis, clone PLACE3000213, weakly similar
 DEFINITION to COMPLEMENT RECEPTOR TYPE 2 PRECURSOR.

ACCESSION AK075234

VERSION AK075234.1 GI:22761190

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 Isogai, T., Ota, T., Nishikawa, T., Hayashi, K., Otsuki, T.,
 Sugiyama, T., Suzuki, Y., Sato, H., Nagai, K., Sugano, S., Ishii, S.,
 Kawai-Hio, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y.,
 Kojima, S., Nagahara, K., Masuho, Y., Ono, T., Okano, K., Yoshikawa, Y.,
 Aotsuka, S., Sasaki, N., Hattori, A., Okumura, K., Iwayanagi, T. and
 Ninomiya, K.
 NEDO human cDNA sequencing project

Unpublished

2 (bases 1 to 4088)

Isogai, T. and Otsuki, T.

Direct Submission

Submitted (25-MAR-2002)

Takao Isogai, Helix Research Institute,
 Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology; cDNA library construction:

COMMENT

Institute of Medical Science, University of Tokyo, Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers

FEATURES

source
1..4088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE3000213"
/tissue_type="placenta"
/clone_lib="PLACE3"
/note="cloning vector: pME18SFL3"
misc_difference 1474..1475
/note="compared to AL158158.8"
/replace="a"
misc_difference 2658..2660
/note="compared to AL158158.8"
/replace="gga"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4088
Score: 6950.00 Matches: 1232
Percent Similarity: 99.20% Conservative: 1
Best Local Similarity: 99.12% Mismatches: 10
Query Match: 34.80% Indels: 1
DB: 9 Gaps: 0

US-09-977-053-4 (1-3571) x AK075234 (1-4088)

QY 2329 LeuGluAsnGlnLeuValLeuLysGluLeuThrThrGluValGlyValValThrPheSer 2348
DB 1 ATGGAAACCCGCTAGTATTAAGGAGGTGACCCAGGTAGGAGTGTGACATTTTCC 60
QY 2349 CysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerGlnGln 2368
DB 61 TGTAAAGAGGCGATGCTCTGACAGGCCCTCTGCTGGAATGCTTGCCATCCCGCAA 120
QY 2369 TrpAsnAspSerPheProValCysLysIleValLeuCysThrProProProLeuIleSer 2388
DB 121 TCGAATGACTTTCTCCCTGTTGTAGATGTTCTTGTATGCCCACTCCCTCAATTTCC 180
QY 2389 PheGlyValProIleProSerSerAlaLeuHisPheGlySerThrValLysThrSerCys 2408
DB 181 TTGGTGTCCTCCATCTCTTCTGCTCTTCATTTTGGAGTACTGTCAAGTATCTTGT 240
QY 2409 ValGlyGlyPheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrp 2428
DB 241 GTAGGTGGGTTTTCTTAGAGGGAATTTCTACCACTCTGCCAATCTGATGGCCTGG 300
QY 2429 SerSerProLeuProGluCysValProValGluCysProGlnProGluGluIleProAsn 2448
DB 301 AGCTCTCCACTGCCAGATGTGTTCCAGTAGAATGTCCCAACTGAGGAAATCCCAAT 360
QY 2449 GlyIleIleAspValGlnGlyLeuAlaThrLeuSerThrAlaLeuThrCysLysPro 2468
DB 361 GGAATCATTTGATGTGACAGCCCTTGCTATCTCAGCACAGCTCTCTATACCTGCAAGCA 420
QY 2469 GlyPheGluValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGly 2488
DB 421 GCCTTTGAATGGTGGGAATACTACCACTTTGTGGAGAAAATGGTCACTGGCTTGA 480
QY 2489 GlyLysProThrCysLysAlaIleGluCysLeuLysProLysGluIleLeuAsnGlyLys 2508
DB 481 GGAACCAACCAATGTAAGCCATTGAGTGCCTGAAACCCAGAGAGATTTGAATGGCAAA 540
QY 2509 PheSerThrThrAspLeuHisThrGlyGlnThrValThrThrSerCysAsnArgGlyPhe 2528
DB 541 TTCTCTTACCGGACCTACACTATGACAGACCGCTTACCTACTCTTTGCACCGAGCTTT 600
QY 2529 ArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAla 2548

DB 601 CGGCTCGAAGTCCCAAGTCCCTTGACCTGTGTAGACACAGGTGATGGGATGTAGATGCC 660
QY 2549 ProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGlyPheValGlu 2568
DB 661 CCATCTTGCAATGCCATCCACTGTGATTCGCCACACCCATTCGAAATGGTTTTGTAGAA 720
QY 2569 GlyAlaAspThrSerThrGlyAlaIleIleThrSerCysPheProGlyPheGlnVal 2588
DB 721 GGTGCAGATTACAGCTATGGTGCATATCATCATCATGTGCTTCCCTGGGTTTTCAGGTG 780
QY 2589 AlaGlyHisAlaMetGlnThrCysGluGlnSerGlyTrpSerSerSerIleProThrCys 2608
DB 781 GCTGCTCATGCCATCGACACCTGTGAGAGTCAAGTGTCAAGTTCCATCCCAACATGT 840
QY 2609 MetProIleAspCysGlyLeuProProHisIleAspPheGlyAspCysThrLysLeuLys 2628
DB 841 ATGCCAATAGACTGTGGCTCCCTCTCATATAGATTTTGGAGACTGTACTAAACTCAA 900
QY 2629 AspAspGlnGlyThrPheGluGlnGluAspMetMetGluValProThrValThrPro 2648
DB 901 GATGACCGGATATTTTGGCAGAGACGACATGATGAGATTTCATATGTGACTCCT 960
QY 2649 HisProProThrHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLysGluSerPro 2668
DB 961 CACCTCTTATCATTTTGGGAGCAGCGGCTAAAACCTGGGAAATACAAAGAGTCTCCT 1020
QY 2669 AlaThrHisSerSerAsnPheLeuThrGlyThrMetValSerThrCysAsnProGly 2688
DB 1021 GCTACACATTCATCAAACTTTCTGTATGTATGATGATGATGATGATGATGATGATG 1080
QY 2689 TyrGluLeuLeuGlyValLeuValLeuLysGlnGluAspGlyThrTrpAsnGlySer 2708
DB 1081 TATGAACTTCTGGGAAACCTGTGTGATCTGCCAGAGATGGAATCTGGAATGCGAGT 1140
QY 2709 AlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsnGlyPheLeu 2728
DB 1141 GCACCATCTCGCAATTCGAATGTGACTTGCCTACTCTCTCAAAATGGCTTTTG 1200
QY 2729 ArgPheThrGluThrSerMetGlySerAlaValGlnThrSerCysLysProGlyHisIle 2748
DB 1201 CTTTTTACAGACACTAGCATGGGAAGTGTGTGAGTATAGCTGTAAACCTGGACACAT 1260
QY 2749 LeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGlyAlaSerPro 2768
DB 1261 CTAGCAGGCTCTGACTTAAAGCTTTGTCTAGAGAATAGAAAGTGGAGTGGTGGCTCCCA 1320
QY 2769 ArgCysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySerIleLysGly 2788
DB 1321 CGCTGTGAAGCCATTTTCATGCAAAAGCCAAATCCAGTCAATGAATGGATCCATCAAGGA 1380
QY 2789 SerAsnThrThrThrLeuSerThrLeuThrCysAspProGlyThrValLeuAsn 2808
DB 1381 AGCAACTACATACCTAGACAGCTGTGTACTATGAGTGTGCCCCCGGATATGTCTGAAT 1440
QY 2809 GlyThrGluArgThrCysGlnAspAspLysAsnTrpAspGluAspGluProIleCys 2828
DB 1441 GGCACCTGAGAGGAGACATGCCAGGATGACA-AACTGGGATGAGGATGAGCCCATTTGC 1499
QY 2829 IleProValAspCysSerSerProValSerAlaAsnGlyGlnValArgGlyAspGlu 2848
DB 1500 ATTCTGTGGAAGTGTGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 1559
QY 2849 TyrThrPheGlnLysGluIleGluThrCysAsnGluGlyPheLeuLeuGluGlyVala 2868
DB 1560 TACACATTCGAAAGAGATTTGAATACATCTGCAATGAAAGGTTCTTCTGTAGGGAGCC 1619
QY 2869 ArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyValaThrProAspCysValPro 2888
DB 1620 AGGAGTGGGTTGTCTTCCCAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1679
QY 2889 ValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGlyLeuAspThrGly 2908
DB 1680 GTCAGATGTGCCACCCCGCCACAACTGGGCAATGGGGTGTGAGGAGGCTGTGATATGCT 1739

```
Qy 2909 PheMetIysGluValThrPheHisCysHisGluClyTyrIleLeuHisGlyAlaProLys 2928
Db 1740 TTCATGAAGGAAGTAAATTCATTCATGTCATGAGGCTACATCTTCACGCGTCTCCAAA 1799
Qy 2929 LeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCysLysProValAsn 2948
Db 1800 CTCACCTGTCACTCAGTCAGATGGCACTGGATGTCAGAGATTCCTCTCTGTAAACCACTCAAC 1859
Qy 2949 CysGlyProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGly 2968
Db 1860 TGTGACCTCTCGAAGATCTTCGCCATGGTTCCTTAATGGTTTCTTTATTCATGGG 1919
Qy 2969 GlyHisIleGlnTyrGlnCysPheProGlyTyrIysLeuHisGlyAsnSerSerArgArg 2988
Db 1920 GGCATATACAGTATCAGTCTTCCTGGTATTAAGCTCCATGAGAAATTCATCAAGAAGG 1979
Qy 2989 CysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCysArgCysSer 3008
Db 1980 TGCCTCTCCAAATGGCTCTGAGTGGGAGCTCACCTTCCTGCTGCTTCGCAGATGTTC 2039
Qy 3009 ThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAla 3028
Db 2040 ACACCACTAATTTGAATATGGAATGTCAATGGGACAGATTTTGACTGTGGAAAGCCAGCC 2099
Qy 3029 ArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAla 3048
Db 2100 CGGATTCAGTGTCTTCAAGGCTTCAAGCTCTTAGACTTTCTGAAATCACTCTGAAAGCC 2159
Qy 3049 AspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuPro 3068
Db 2160 GATGGCCAGTGGAGCTCTGGGTTCCTCCACTGTGACACACACTTCTGTGGTCTCTTCCA 2219
Qy 3069 MetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGluAsnValIleThrTyr 3088
Db 2220 ATGATACCAATATGCTTCATCAGTGAGACCACTCTTGGAAAGAAAATGTGTAATCTTAC 2279
Qy 3089 SerCysArgSerGlyTyrValIleGlnClySerSerAspLeuIleCysThrGluLysGly 3108
Db 2280 AGCTGCAGGCTCTGATATGTATATACAGGCGATTCAGATCTGAGTGTGACAGAGAAAGG 2339
Qy 3109 ValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerVal 3128
Db 2340 GTATGGAGCCAGCTTATTCAGTCTGTGAGCCCTTGCTGCTGCTGCCACCGTCTGTC 2399
Qy 3129 AlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCys 3148
Db 2400 GCCAATGCACTGGAGAGGCAACACCTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2459
Qy 3149 LeuGluGlyTyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyArg 3168
Db 2460 CTGAAGGTTATATGATGATACAGATACAGATACATCACTGTGAGAAAGATGGTGC 2519
Qy 3169 TrpPheProGluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleThr 3188
Db 2520 TGGTTCCTCGAGAGAACTCTCGAGTCTTAAATAATGCTCTCTCCGGGAAAACATAACA 2579
Qy 3189 HisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAla 3208
Db 2580 CATATACTTGTATACATGGGAGCAATTCAGTGTGAATAGGCAAGTTTCTGTGTATGTGCA 2639
Qy 3209 GluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGlu 3228
Db 2640 GAAGGATATATCTTTGAGTGAGTTAACAATATACATATCAGTATGTACAGTTGATGGAACTGGAG 2699
Qy 3229 ProProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGlu 3248
Db 2700 CCACCATCTCCGATGATCTTTCAGTCCAGTCTGAGTCTTGGGAAACCTGAAAGTCCAGAA 2759
Qy 3249 HisGlyPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGlu 3268
Db 2760 CATGGATTTGTGGTGGCAGTAAATACACCTTTGAAAGCAACAATTTATTTATTCAGTGTGAG 2819
```

```
Qy 3269 ProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSer 3288
Db 2820 CTTGGCTATGAACTAGAGGGGACAGAGGAGCTGTCTGCGCAGGAGAACAGACAGCTGGAGT 2879
Qy 3289 GlyGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGly 3308
Db 2880 GGAGGGGTGGCAATATGCAAGACAGCAGGTGTGAACTCCACTTGAATTTCTCAATGGG 2939
Qy 3309 LysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGly 3328
Db 2940 AAAGCTGACATTTGAAACAGAGACCTGGACCAACCTGGTATATTTCTGCAACAGAGGC 2999
Qy 3329 TyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisPro 3348
Db 3000 TACAGTCTTGGAGGGCCATCTGAGGCACACTGCACAGAAATGGAACTGGAGCCACCCA 3059
Qy 3349 ValProLeuCysLysProAsnProCysProValProPheValIleProGluAsnAlaLeu 3368
Db 3060 GTCCCTCTCTGCAAAACCAATCCATGCCCTGCTTCTTTGTGATTTCCCGAGAATGCTCTG 3119
Qy 3369 LeuSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPhe 3388
Db 3120 CTGTCTGAAAGGAGTTTATGTTGATCAGAAATGTGTCATCAATGTAGGGAAGGTTTT 3179
Qy 3389 LeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSer 3408
Db 3180 CTGCTGCAAGGCCACCGGATCATTTACCTGCAACCCCGACAGAGACGTGACACAGACAGC 3239
Qy 3409 AlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArg 3428
Db 3240 GCCAATGTGAAAAAATCTCATGTGTCACAGCTCACGTAGAAAAATGCAATTTGCTCGA 3299
Qy 3429 GlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeu 3448
Db 3300 GGCCTACATTTATCAATATGGAGACATGATCCTACTCATGTATTACAGTGGATACATGTTG 3359
Qy 3449 GluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProIleCys 3468
Db 3360 GAGGTTTCCAGAGGAGTGTGTTGTTAGAAATGGAACATGGACATCCTCTCTATTTCG 3419
Qy 3469 ArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCys 3488
Db 3420 AGAGCTGTCTGTCGATTTCCATGTGCAATGCGGGGCATCTGCAACACGCCCAATGCTGT 3479
Qy 3489 SerCysProGluGlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeuProCys 3508
Db 3480 TCCGTGTCAGAGGGCTGGATGGGGCGCTCTGTGAAGAAACCAATCAGCAATTTCTCCCTGT 3539
Qy 3509 LeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGlyTrpThrGly 3528
Db 3540 CTGAACGGAGGTCTGCTGTGTGCCCCCTTACCAGTGTGACTGCCCGCTGGCTGGAGCGGG 3599
Qy 3529 SerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArg 3548
Db 3600 TCTCGCTGTCATGCACTGTTTGGCAGTCTCCCTGCTTAAATGGTGGAAAAATGTGTAGA 3659
Qy 3549 ProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArg 3568
Db 3660 CCAAAACCGATGTCATGCTTCTTCTTCTTGGAGCGGACATACTGTTCAGGAAAGAGGG 3719
Qy 3569 ThrGlyPhe 3571
Db 3720 ACTGGGTTT 3728
```

```
RESULT 19
HSM803724
LOCUS HSM803724 3914 bp mRNA linear PRI 10-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKF2p6670I.713 (from clone DKF2p6670I.713).
ACCESSION AL832416
VERSION AL832416.1 GI:21732980
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
```

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Rutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 3914)
TITLE	Koehler,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.
COMMENT	Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY Clome from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

```

FEATURES
  source
    i. .3314
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="RZPD:DKFZp667O1713"
      /db_xref="taxon:9606"
      /clone="DKFZp667O1713"
      /tissue_type="lymph node"
      /clone_lib="667 (synonym: hlmo2) . Vector pSport1; host
DH10B; sites NotI + SalI"
      /dev_stage="adult"
      3773..3778
      polyA_signal
      polyA_site
      3797

```

US-09-977-053-4 (1-3571) x HSM803724 (1-3914)	
QY	2393 IleProSerSerAlaLeuHisPheGlySerThrValIysTyrSerCysValGlyGlyPhe 2412
DB	1 ATTCTCTCTCTGCTCTTCAATTTGGAGTACTGTCAAGTATTTCTGTGTAGTGGTTTT 60
QY	2413 PheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeu 2432
DB	61 TTCTTAAGAGCAAAATCTACCACTCTGCCAACCTGATGCCACTGGAGCTCTCCACTG 120
QY	2433 ProGluCysValProValGluCysProGlnProGluGluIleProAsnGlyIleLeuAsp 2452
DB	121 CCAGNATGTGTTCCAGTAGAATGTGCCAACCTTGAGGAAATCCCCAATGGAATCATGTAT 180
QY	2453 ValGlnGlyLeuAlaTyrLeuSerThrAlaLeuTyrThrCysIysProGlyPheGluLeu 2472
DB	181 GTGCNAGGCTTGGCTATCTCAGCACAGCTCTCTATACCTGCAGAGCCAGCGCTTGAATTG 240
QY	2473 ValGlyAsnThrThrThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyIysProThr 2492
DB	241 GTGGGAAATACTACCACTCTTGTGGAGAAAATGGTCACTGGCTTGAGGAGAAACCAACA 300
QY	2493 CysIysAlaIleGluCysLeuIysProGlyGluIleLeuAsnGlyIysPheSerTyrThr 2512
DB	301 TGTAAAGCCATTGAGTGCCTGAAACCCAGAGAGATTTTGAATGGCAAAATCTCTTACACG 360
QY	2513 AspLeuHisTyrGlyGlnThrValThrTyrSerCysAsnArgGlyPheArgLeuGluGly 2532
DB	361 GACCTACATATGACACAGCCGTACTCTCTTGCAACCGAGGCTTTCGGCTTCGAGGT 420
QY	2533 ProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsn 2552
DB	421 CCCAGTGCCTTGACCTGTTTAGAGACAGGTGATTGGATGTAGATGCCCACTTCGCAAT 480

1561 GTAACTTCACCTGTCATGAGGGCTACATCTTGACCGGTGCTCCAAACTCCTGTCAG 1620
2933 SerAspGlyAsnTrpAspAlaGluLeuProLeuCysLysProValAsnCysGlyProPro 2952
1621 TCAGATGGCAACTGGGATGACAGAGATTCCTCTCTGTAACACAGTCAACTGTGGACCTCT 1680
2953 GluAspLeuAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGln 2972
1681 GAAGATCTTGCCCATGGTTTCCCTTAATGGTGTTCCTTTATTCATGGGGGCCATATACAG 1740
2973 TyrGlnCysPheProGlyTyrLysLeuHisGlyAsnSerSerArgArgCysLeuSerAsn 2992
1741 TATCAGTGTCTTCTGGTATAGCTCCATGGAAATTCATCAAGAAGTGGCTCTCCAAAT 1800
2993 GlySerTrpSerGlySerProSerCysLeuProCysArgCysSerThrProValIle 3012
1801 GGCTCTCGAGTGGCAGCTCACCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
3013 GluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGlnCys 3032
1861 GAATATGGAACTGTCAATGGAGACAGATTTTGACTGTGGAAAGGAGCCCGGATTCAGTGC 1920
3033 PheLysGlyPheLysLeuLeuGlyLysSerGluIleThrCysGluAlaAspGlyGlnTrp 3052
1921 TTCAGAGGCTTCAGCTCCTAGGACTTCTGAAATCACCTGTGAAGCCGATGCCAGTGG 1980
3053 SerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIleProAsn 3072
1981 AGCTCTGGGTTCCTCCCACTGCTGAACACACTTCTGTGTGTGTCTCTTCCTCCATGATACCAAT 2040
3073 AlaPheIleSerGlnThrSerSerTrpLysGluAsnValIleThrTyrSerCysArgSer 3092
2041 GCCTTCATCAGTGGACCACTCTTGGAGGAAATGTGATACTTACAGCTGCGAGTCT 2100
3093 GlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyValTrpSerGln 3112
2101 GGATATGTATACAGGACAGTTCAGATCTCATTTGTACAGAGAAAGGGGTATGGAGCCAG 2160
3113 ProTyrProValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaVal 3132
2161 CCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCTCCCACTGCTGTGTGCGCAATGCAATG 2220
3133 AlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLeuGluGlyTyr 3152
2221 GCAACTGGAGAGGACACACCTATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2280
3153 ThrMetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGlu 3172
2281 ACGATGGATACAGATACAGATACATTCACCTGTGAGAAAGATGGTGGTGGTTCCTGAG 2340
3173 ArgIleSerCysSerProLysCysProLeuProGluAsnIleThrHisIleLeuVal 3192
2341 AGAATCTCTGCACTCTTAAATAATGCTCTCTCCCGGAAACATACACATATATCTGTT 2400
3193 HisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThr 3212
2401 CATGGGAGACATTTCAAGTGTGAATAGCAAGTTCTGTGTGTGTGTGTGTGTGTGTGTGT 2460
3213 PheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSer 3232
2461 TTTGAGGGAGTTAAACATATCATATGATGATGATGATGATGATGATGATGATGATGATG 2520
3233 AspGluSerCysSerProValSerCysGlyLysProGluSerProGluHisGlyPheVal 3252
2521 GATGAATCTTTCAGTCCAGTCTCTTGTGGGAAACCTGAAAGTCCAGAAATGATGTTGTG 2580
3253 ValGlySerLysThrThrPheGluSerThrIleIleTyrGlnCysGluProGlyTyrGlu 3272
2581 GTTGGCAGTAAATACACCTTTGAAGACCAATTTATTTATCAGTGTGAGCTGCTGTGAA 2640
3273 LeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyGlyValAla 3292

2641 CTAGAGGGGAACAGGGAAACGTGCTCTGCCAGGAGAACACAGACAGTGGAGTGGGGTGGCA 2700
3293 IleCysLysGluThrArgCysGlyThrProLeuGluPheLeuAsnGlyLysAlaAspIle 3312
2701 ATATGCAAGAGAGACAGGTGTGAACCTCCACTTCAATTTCTCAATGGGAAAGCTGACAT 2760
3313 GluAsnArgThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGlu 3332
2761 GAAACAGAGAGAGCTGAGCCACAGTGTATATTCCTGCAACAGAGGCTACAGTCTTGAA 2820
3333 GlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCys 3352
2821 GGGCCATCTGAGGCACACTGCACAGAAAATGGAACCTGGAGCCACCCAGTCCCTCTCTGC 2880
3353 LysProAsnProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLys 3372
2881 AAACCAATCCATGCTCTGCTTTGTGATTCCTGAGATGCTCTGCTGCTGCTGCTGCTG 2940
3373 GluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGly 3392
2941 GAGTTTATGTTGATCAGATGTGTCATCAATGTAGGAGAGTTTCTGCTGCGAGGC 3000
3393 HisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGlu 3412
3001 CACGGCATCATTAACCTGCAACCCCGACGAGAGCTGGACACAGACAGCCCAATGTGAA 3060
3413 LysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyr 3432
3061 AAAATCTCATGTGTGTCACAGCTCAGCTAGAAAATGCAATTTGCTGAGGGGTACATAT 3120
3433 GlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGluGlyPheLeu 3452
3121 CAATATGGAGACATGATCCTACTCATGTACAGTGTGATACATGTTGGAGGGTTTCTG 3180
3453 ArgSerValCysLeuGluAsnGlyThrTrpThrSerProProIleCysArgAlaValCys 3472
3181 AGCAGTGTGTTGTTAGAAAATGGAACATGACATCACCTCTCTATTTGACAGAGTGTCTGT 3240
3473 ArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGlu 3492
3241 CCAATTTCCATGTAGAAATGGGGCATCTGCCAACGCCCAATGCTGTGCTGCTGACAG 3300
3493 GlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGly 3512
3301 GGTGAGTGGGGGCGCTCTGTGAGAACCAATCTGCATTTCTTCTGCTGCAACGGAGGT 3360
3513 ArgCysValAlaProTyrGlnCysAspCysProProGlyTrpThrGlySerArgCysHis 3532
3361 CGCTGTGTGGCCCTTACCAAGTGTGACTGCGCCGCTGGAGCGGGTCTCGCTGTCTAT 3420
3533 ThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCys 3552
3421 ACAGCTGTGTCAGTCTCCCTGCTTAAATGGTGGAAAATGTGTAGACCAACCGATGT 3480
3553 HisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGlyPhe 3571
3481 CACTGCTTCTTCTTGGACGGGACATACATGTTCCAGGAAAAGAGGAGGACTGGGTTT 3537

RESULT 20

AL158158/c

LOCUS

DEFINITION

complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AL158158 194835 bp DNA linear PRI 05-APR-2001
Human DNA sequence from clone RP11-427L11 on chromosome 9q31.2-32,
complete sequence.

AL158158

AL158158.14 GI:13559997

HTG.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 194835)

Williams, S.

Direct Submission

Submitted (05-APR-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Apr 6, 2001 this sequence version replaced gi:12733508.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr9>

RP11-427L11 is from the library RPCL11-2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

This sequence is the entire insert of clone RP11-427L11.

repeat_region	/note="L1M9A repeat: matches 6027. .6305 of consensus" 5871. .5938
repeat_region	/note="AluJ repeat: matches 1. .68 of consensus" 5967. .6648
repeat_region	/note="L2 repeat: matches 1284. .2090 of consensus" 6549. .6948
repeat_region	/note="AluSg repeat: matches 1. .300 of consensus" 6949. .7003
repeat_region	/note="L2 repeat: matches 2090. .2136 of consensus" 7072. .7482
repeat_region	/note="L2 repeat: matches 2295. .2695 of consensus" 7588. .7789
misc_feature	/note="MIR repeat: matches 12. .236 of consensus" complement (8285. .8469)
misc_feature	/note="match: STS: Em:G05863" complement (8329. .8496)
misc_feature	/note="match: GSS: Em:AQ227987" complement (8329. .8495)
misc_feature	/note="match: GSS: Em:B65750" complement (8339. .8502)
misc_feature	/note="match: GSS: Em:AQ110861" 8354. .8529
misc_feature	/note="match: STS: Em:G54088" join(8370. .8532,9091.9156,15133. .15192,15671. .15768)
repeat_region	/note="match: STS: Em:G29952" 8591. .8662
repeat_region	/note="36 copies 2 mer ta 81% conserved" 9732. .9893
misc_feature	/note="AluJ repeat: matches 130. .287 of consensus" complement (9737. .9894)
misc_feature	/note="match: GSS: Em:AQ281101" complement (9886. .10119)
misc_feature	/note="match: GSS: Em:B53125" complement (9897. .10091)
misc_feature	/note="match: STS: Em:HSAL08WE1" complement (9907. .10123)
repeat_region	/note="match: GSS: Em:B17453" 9932. .9967
repeat_region	/note="18 copies 2 mer ta 100% conserved" 9973. .10105
repeat_region	/note="AluJ repeat: matches 1. .133 of consensus" 10133. .10392
misc_feature	/note="AluSg repeat: matches 38. .299 of consensus" 10183. .10319
repeat_region	/note="match: GSS: Em:AQ782105" 10404. .10492
misc_feature	/note="WERSB repeat: matches 7. .97 of consensus" 10992. .11078
repeat_region	/note="match: GSS: Em:AQ544271" 11354. .11653
repeat_region	/note="AluSg repeat: matches 1. .300 of consensus" 11742. .12929
repeat_region	/note="Tigger1b repeat: matches 2. .1231 of consensus" 12987. .13146
repeat_region	/note="MIR repeat: matches 89. .248 of consensus" 13153. .13468
repeat_region	/note="AluSg repeat: matches 1. .312 of consensus" 13678. .13848
repeat_region	/note="L2 repeat: matches 2537. .2703 of consensus" 14394. .14609
repeat_region	/note="MIR repeat: matches 12. .257 of consensus" 15816. .15902
misc_feature	/note="match: GSS: Em:B46502" 15900. .15947
repeat_region	/note="24 copies 2 mer aa 79% conserved" 16257. .16558
repeat_region	/note="AluSg repeat: matches 1. .302 of consensus" 16956. .17220
repeat_region	/note="AluSg repeat: matches 38. .304 of consensus" 17415. .17488
repeat_region	/note="37 copies 2 mer aa 71% conserved" 18552. .18845
repeat_region	/note="AluSg repeat: matches 1. .293 of consensus" 18552. .18845

```

repeat_region 19629..19689
misc_feature /note="L2 repeat: matches 2685..2743 of consensus"
repeat_region 19779..21330
misc_feature /note="CpG Island"
repeat_region 21481..21679
misc_feature /evidence=not experimental
repeat_region 22591..22818
misc_feature /note="MIR repeat: matches 48..262 of consensus"
repeat_region 22591..22818
misc_feature /note="MIR repeat: matches 23..258 of consensus"
misc_feature complement (23067..23458)
misc_feature /notes="match: GSS: Em:B43385"
misc_feature complement (23074..23458)
misc_feature /notes="match: GSS: Em:B43132"
repeat_region 23731..23893
misc_feature /note="MER5A repeat: matches 17..180 of consensus"
repeat_region 24008..24365
misc_feature /note="match: GSS: Em:AQ088039"
repeat_region 24107..24196
misc_feature /note="MIR repeat: matches 105..202 of consensus"
repeat_region 24953..25091
misc_feature /note="L1M8 repeat: matches 6145..6287 of consensus"
repeat_region 25184..25589
misc_feature /note="MER50 repeat: matches 1..436 of consensus"
repeat_region 25620..25718
misc_feature /note="MER50 repeat: matches 640..734 of consensus"
repeat_region 25839..25991
misc_feature /note="L1M1 repeat: matches 5212..5366 of consensus"
repeat_region 25992..26078
misc_feature /note="AluSp/q repeat: matches 174..260 of consensus"
repeat_region 26084..26798
misc_feature /note="L1M1 repeat: matches 5347..6036 of consensus"
repeat_region 27326..27430
misc_feature /note="L1M6 repeat: matches 4066..4176 of consensus"
repeat_region 27431..27740
misc_feature /note="AluX repeat: matches 1..310 of consensus"
repeat_region 27741..28086
misc_feature /note="L1M6 repeat: matches 4176..4455 of consensus"
repeat_region 28087..28226
misc_feature /note="AluJb repeat: matches 3..132 of consensus"
repeat_region 28227..28530
misc_feature /note="AluX repeat: matches 1..297 of consensus"

Alignment Scores:
Pred. No.: 0 Length: 194835
Score: 6218.00 Matches: 1228
Percent Similarity: 49.74% Conservative: 1
Best Local Similarity: 49.70% Mismatches: 2
Query Match: 31.13% Indels: 1242
DB: 9 Gaps: 2

US-09-977-053-4 (1-3571) x AL158158 (1-194835)
Qy 1993 TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSer 2012
Db 176047 TATACTTGTGCTGTGTGACCAATGTAATGCTGCGCCGACGCAAGTGGAGTAGAAGT 175988
Qy 2013 AspGlnGlnCysLeuAlaValSerCysAspGluProIleValAspHisAlaSerPro 2032
Db 175987 GACCAGCAGTCCGCTGCTCTCTGTGATGAGCCACCATTGTGGACCGCTCTCCA 175928
Qy 2033 GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSer 2052
Db 175927 GAGCTAGCCCATCGGCTCTTGGAGACATTCATTCTTACTACTGCTCTGATGGTTACAGC 175868
Qy 2053 LeuAlaAspAsnSerGlnLeuLeuCysAspAlaGlnGlyLysTrpValProProGluGly 2072
Db 175867 CTAGCAGACAAATCCCGACTTCTCTGCAATGCCAGGCGCAAGTGGGTACCCCGAGAGGT 175808
Qy 2073 GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer 2092
Db 175807 CAAGCATGCCCGTGTATAGTCTCATTTCTGTGAAMAACCTCCATCGGTTTCTATAGC 175748
Qy 2093 IleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys 2112

```

```

Db 175747 ATCTTGGAATCTGTGAGCAAAAGCAAAATTTGCAGCTGCTCAGTTGTGAGCTTTAAATGC 175688
Qy 2113 MetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrp 2132
Db 175687 ATGGAAGGCTTTGTACTGAACACCTCAGCAAAAGATTAATGTATGAGAGGTTGGGCGAGTG 175628
Qy 2133 AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle 2152
Db 175627 AACCTTCCCCCATGTTCATCCAGTGCATCCCTGTGCGGTGTGGAGAGCCACCAAGCATC 175568
Qy 2153 MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys 2172
Db 175567 ATGAATGGCTATGCAAGTGCATCAAACTACAGTTTTCGAGCCATGGTGGCTTTACAGCTGC 175508
Qy 2173 AsnLysGlyPheTyrIleLysGlyGluLysLysSerThrCysGluAlaThrGlyGlnTrp 2192
Db 175507 AACAGGGGTTTACATCAACGAGGGAAGAAGAGAGACCTCCGAGGCCACAGGGCGAGTG 175448
Qy 2193 SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn 2212
Db 175447 AGTAGTCTTATACGACGTGCCACCCCGTATCTTGTGTGAACCACTTAAGTTGAGAT 175388
Qy 2213 GlyPheLeuGlu----- 2216
Db 175387 GGCTTTCTGGA-GGTAAGAGACCAATTAGCAAAATGCTGTGGCTTTTGTACAGGGGCC 175329
Qy 2216 ----- 2216
Db 175328 ATAATAAATAAATAAACAACAATAATTTGAGTCCAAACACTGCCAAGTTCACTAAGGC 175269
Qy 2216 ----- 2216
Db 175268 ATGGAACCTGGATTTACTATTATCACTAACAAAGATCATGATAATCTTTTACATTT 175209
Qy 2216 ----- 2216
Db 175208 TCATAATTAATTCATCTCTTGATGAATAATTCATCTGCTAAGTGTGTAATCATCTTC 175149
Qy 2216 ----- 2216
Db 175148 TGGAGGTAAGAGACTAATTAGCAAAATGCTGTGGCTTTTCATCACTTACAGGGCCAGGCC 175089
Qy 2216 ----- 2216
Db 175088 ACTGTGATGAACCCACTATTTCAGAGCGGGCAATGAACAAATTAATTTATGAAAAACAG 175029
Qy 2216 ----- 2216
Db 175028 AAATGGCTAAAAATCTACAGGTAGTGAAGACTGGAAAAAATAATTAATGAAGATGATTGA 174969
Qy 2216 ----- 2216
Db 174968 ATCTTTAGTATATAAATTTCTGCAATTTTACTAGTGTATAGGCATTTTCAAGTCTTTGG 174909
Qy 2216 ----- 2216
Db 174908 CAACAGCTTAGCCCTAAATAGGAGTTAGTAATAACCAATTTTAAATAATAATTTTGT 174849
Qy 2216 ----- 2216
Db 174848 TACAGAGGTTTTTCAGGATGATATATGCTTTTATTTATAGAAGCAAGGTCTAAATGA 174789
Qy 2216 ----- 2216
Db 174788 ACCTAGAGAAATATATATTAATTTTAAATGCTTAGACAAAGATAGGTAACCAATTA 174729
Qy 2216 ----- 2216
Db 174728 TGTAGTTTTTACAATCTGTTTTTGGTAAACCTAAAGGTAAATTTTCACTGTGATGGACAAA 174669
Qy 2216 ----- 2216

```



```
Db 170289 ATAAATCGATACCTTTTTCAGACATGTTTCTTAATATAATTATCATGATCATATTTCTG 170230
QY 3146 ----- 3146
Db 170229 TTGAGCTTCGTGAACAAACAGCAAGATCCTCTAGCTAGCTTGAAAGCTCGGAATCGT 170170
QY 3146 ----- 3146
Db 170169 GAAGCTGTTTTTAAACATCAGGATAGTAAATTTCTTTATGATTAATACTTTGTGATTA 170110
QY 3146 ----- 3146
Db 170109 ATTGACATGCAATATTTGGGACATTAAGGCAAAAACCTGAAGAATTCAGGCCATGATC 170050
QY 3146 ----- 3146
Db 170049 ATGCCAAATATCAGGATTTAAAGAAATATCTTATGTAAAGGACCTAGAGCAGCACTTGG 169990
QY 3146 ----- 3146
Db 169989 CACAAAGTAGGTATTCATAATACATANGAATTATGCTGTATTCCTACTAGTATCTGAGTC 169930
QY 3146 ----- 3146
Db 169929 CATGTAGGTATTCATAATACATATGAATTAATGCTGTATTCCTACTAGTATCTGAGTCCAC 169870
QY 3146 ----- 3146
Db 169869 GTAGGTATTCATAATACATATTAATTAATGCTGTATTCCTACTAGTATCTGAGTCCACGTA 169810
QY 3146 ----- 3146
Db 169809 GGTATTCATAATACATATGAATTAATGCTGTATTCCTACTAGTATCTGAGTCCACGTAAGT 169750
QY 3146 ----- 3146
Db 169749 ATTCAATACAAATGAATTAATGCTGTATTCCTACTAGTATCTGAGTCCATGTAGTATT 169690
QY 3146 ----- 3146
Db 169689 CATAATACATATGAATTAATGCTGTATTCCTACTAGTATCTGAGTCCATGTAGTATTAT 169630
QY 3146 ----- 3146
Db 169629 AATACATATGAATTAATGCTGTATTCCTACTAGTATCTGAGTCCATGTAGTATTATTAAT 169570
QY 3146 ----- 3146
Db 169569 ACATATTAATTAATGCTGTATTCCTACTAGTATCTGAGTCCATTTGGAGCTGTAGAAAT 169510
QY 3146 ----- 3146
Db 169509 TCCTGAAGCTGTAGCTGTTTATAGATAAAAAATGAATTCATAATAATAATATCCAAAT 169450
QY 3146 ----- 3146
Db 169449 AAGGTAATAATACATAAAACACTTAAAGTAATTTTGTAGTCTTGATTAACCTTTTCATGTGG 169390
QY 3146 ----- 3146
Db 169389 ATAACTGAGAACTGATCTGGTTATTAACCGTTTTTTTGTGTTTTTTTGGAGACAACG 169330
QY 3146 ----- 3146
Db 169329 TCTTGCTGTGTCACCCAGGCTGGAGTGCAGTGGCAGATGTTCTTGGCTCACTGCAGCCTCAA 169270
QY 3146 ----- 3146
Db 169269 CCTCCTTAGTTCAGCAGTCTCCACCCTCTGCTCCCAAGTAGCTGAGACTACAGGCAT 169210
QY 3146 ----- 3146
Db 169209 GTGCCACCGTCCAGCTAAATTTTGTATTTTATAGATGAAGTCTCACTATGTTT 169150
```

```
QY 3146 ----- 3146
Db 169149 CCCAGCTAGTCTCAAGCTCTCTGGACTCAAGTGGTCTCTCTGCTCAGCCTTCCAAAGTG 169090
QY 3146 ----- 3146
Db 169089 CAGGANTTACAGGCATGAGCCACCGCACCCAGCTCACTGGTTGTTATTGTTATGTTGGT 169030
QY 3146 ----- 3146
Db 169029 GTTTTTCATGCAATGCCATCGCTGACATATGATTGTTGTGCAACAATCAATATGTTGC 168970
QY 3146 ----- 3146
Db 168969 ACAAATATCTACTGAGCCTTGTACAGCTTGTCTAAATCCAAGGATAGATCAGAAACTAT 168910
QY 3147 ----- 3152
Db 168909 ATTGTACTTCATTTAAGGCACCTATTTTTCATTTCCATTTGCGAGATGCTCGAAGGTTAT 168850
QY 3153 ThrMetAspThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGlu 3172
Db 168849 ACATGATACAGATACAGATACATTTTCCATCTGTCAGAAAGATGTCGCTGTTCCCTGAG 168790
QY 3173 ArgIleSerCysSerProLysCysProLeuProGluAsnIleThrHisIleLeuVal 3192
Db 168789 AGAATCTCTCGAGTCTTAAAAAATGTCCTCTCCCGAAACATAACACATATCTTGT 168730
QY 3193 HisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGluGlyValThr 3212
Db 168729 CATGGGAGCGATTTTCAGTGTGAATAGGCAAGTTTCTGTCTGTCAGGAGGATATACC 168670
QY 3213 PheGluGlyValAsnIleSerValCysGlnLeu 3223
Db 168669 TTTGAGGAGTTAAACATATCATGATGTCAGGTA 168637

RESULT 21
AX050034
LOCUS AX050034 3706 bp DNA linear PAT 12-JAN-2001
DEFINITION Sequence 47 from Patent WO0071710.
ACCESSION AX050034
VERSION AX050034.1 GI:12226407
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Denefle,P., Rosier-Montus,M.P., Arnould-Reguigne,I., Prades,C. and Clepet,C.
TITLE Expression products of genes involved in diseases related to cholesterol metabolism
JOURNAL Patent: WO 0071710-A 47 30-NOV-2000;
Aventis Pharma S.A. (FR)
FEATURES
source
1..3706
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3706
Score: 6044.00 Matches: 1063
Percent Similarity: 99.72% Conservative: 0
Best Local Similarity: 99.72% Mismatches: 3
Query Match: 30.26% Indels: 0
DB: 6 Gaps: 0

US-09-977-053-4 (1-3571) x AX050034 (1-3706)
```

QY 2506 AsnGlyLysPheSerTrpThrAspLeuHisTyrglyGlnThrValThrTySerCysAsn 2525

2 ATGGCAATCTCTTACAGGACCTACATATGGACAGACCGTTACTTCTTGCAC 61
2526 ArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAsp 2545
62 CGAGGCTTTCCGNCGAGAGGTCAGTGCCTTGACCTGTTTAGAGACAGGTGATGGAT 121
2546 ValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGly 2565
122 GTAGATGCCCCATCTTGCATATGCCATCCATCTGTGATTTCCCAACCCATTTGAATGGT 181
2566 PheValGluGlyAlaAspTrpSerGlyValAlaIleIleTrpSerCysPheProGly 2585
182 TTGTAGAGAGTGCAGATACATATGATGTCATATATCTACATGTCCTCCCTGGG 241
2586 PheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTrpSerSerIle 2605
242 TTTTCAGTGGCTGGTCAATGCCATGCAGACCTGTGAAGAGTCAGGATGCTCATC 301
2606 ProThrCysMetProIleAspCysGlyLeuProProHisIleAspPheGlyAspCysThr 2625
302 CCAACATGATATGCCAATAGACTGTGGCTCCCTCCCTCATATAGATTTTGGAGACTGTACT 361
2626 LysLeuLysAspAspGlnGlyTrpPheGluGlnGluAspAspMetMetGluValProTrp 2645
362 AAATCAGAGATGACAGGATATTTTGACAGAGAGACGACATGATGGAGTTCCATAT 421
2646 ValThrProHisProProTrpHisLeuGlyAlaValAlaLysThrTrpGluAsnThrLys 2665
422 GTGACTCCTCACCTCTCTTATATCTTGGAGAGCTGGCTAAACCTGGGAAATATACAAAG 481
2666 GluSerProAlaThrHisSerSerAsnPheLeuTrpGlyThrMetValSerTrpThrCys 2685
482 GAGTCTCTGCTACACATTCATCAATCTTCTGTATGGTACCATCGTTTTCATACCTGT 541
2686 AsnProGlyTrpGluLeuLeuGlyAsnProValLeuIleCysGlnGluAspGlyThrTrp 2705
542 AANCCAGATATGACTTCTGGGAACTCTGTGCTGATCTCCAGGAGAGATGGACTGG 601
2706 AsnGlySerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGluAsn 2725
602 AATGGCAGTCACCATCTCGCATTTCAATTTGAATGTGACTTGGCTACTGCTCCTGAAAT 661
2726 GlyPheLeuArgPheThrGluThrSerMetGlySerAlaValGlnTrpSerCysLysPro 2745
662 GGCCTTTTGGGTTTATCAGAGACTAGCATGGAGAGTCTGTGAGTATAGCTGTAAACCT 721
2746 GlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSerGly 2765
722 GGACATCTTAGGGGCTCTGACTTAAGGCTTTGTCTAGAGATAGAAAGTGGAGTGGT 781
2766 AlaSerProArgCysGluAlaIleSerCysLysLysProAsnProValMetAsnGlySer 2785
782 GCCTCCCCAGCTGTGAAGCAATTTATGCAAAAAGCCAAATCCAGTCATGAATGGATCC 841
2786 IleLysGlySerAsnTrpThrTrpLeuSerThrLeuTrpTrpGluCysAspProGlyTrp 2805
842 ATCAAGAGAGCAACTACATACCTGAGCAGCTTTGATGATGATGACCCCGGATAT 901
2806 ValLeuAsnGlyThrGluArgArgTrpCysGlnAspAspLysAsnTrpAspGluAspGlu 2825
902 GTGCTGAATGGCACTGAGAGAGAGAACATGCCAGGATGACAAAACCTGGGATGAGATGAG 961
2826 ProIleCysIleProValAspCysSerSerProProValSerAlaAsnGlyGlnValArg 2845
962 CCCATTTGCACTCTGTGGAGCTGAGTTTACCCCCAGTCTCAGCCAAATGGCCAGGTGAGA 1021
2846 GlyAspGluTrpThrPheGlnLysGluIleGluTrpThrCysAsnGluGlyPheLeuLeu 2865
1022 GGAGAGCAGTACACATTTCCAAAAGAGATTTGAATACACTTGGCAATGAAGGGTCTTGCTT 1081
2866 GluCluValArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAlaThrProAsp 2885

1082 GAGGGAGCCAGGAGTCCGGTTTGTCTTTCCTTGCCTTGCCTTGCAGTGGAGCCACTCCCGAC 1141
2886 CysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluCluLys 2905
1142 TGTGTGCTTGTAGATGTCACCCGCCCACTTGGCCAAATGGGTGACGGAAGGCGCTG 1201
2906 AspTrpGlyPheMetLysGluValThrPheHisCysHisGluGlyTrpIleLeuHisGly 2925
1202 GACTATGGCTTCATGAAGAGAGTAACATTCACCTGTCTAGAGGCTACATCTTGACCGT 1261
2926 AlaProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuLys 2945
1262 GCTCCAAAATCTACCTGTAGTCAGTGCAGATGGCACTGGGATGAGAGATTCCTCTCTGTAAA 1321
2946 ProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProAsnGlyPheSerPhe 2965
1322 CCAAGTCAACTGTGGACCTCTGAAAGATCTTGGCCATGGTTTCCCTTAATGGTTTTCCTTT 1381
2966 IleHisGlyGlyHisIleGlnTrpGlnCysPheProGlyTrpLysLeuHisGlyAsnSer 2985
1382 ATTCAATGGGGCCCATATACAGTATCAGTGTCTTCTTCTGGTATATAGCTCATGGAAATCA 1441
2986 SerArgArgCysLeuSerAsnGlySerTrpSerGlySerSerProSerCysLeuProCys 3005
1442 TCAAGAGGTGCTCTCCATGGCTCTGGAGTGGCAGCTCACCTTCTCTGCTGCTCTTGC 1501
3006 ArgCysSerThrProValIleGluTrpGlyThrValAsnGlyThrAspPheAspCysGly 3025
1502 AGATGTTTCCACACCCAGTAATATGAATATGGAACTGTCAATGGGACAGATTTTGACTGTGA 1561
3026 LysAlaAlaArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLysSerGluIleThr 3045
1562 AAGCAGCCCGGATTCAGTGTCTTCAAGGCTTCAAGCTCTTAGGACTTCTGAAATCACC 1621
3046 CysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGly 3065
1622 TGTGAAGCCGATGGCCAGTGGAGCTCTGGGTTCCTCCCACTGTGAACACACTTCTTGTGT 1681
3066 SerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrpLysGluAsnVal 3085
1682 TCTCTTCCATGATGATACCAATGCGTTTCATAGTGGAGCCAGCTCTTGGAGGAAATGTG 1741
3086 IleThrTrpSerCysArgSerGlyTrpValIleGlnGlySerSerAspLeuIleCysThr 3105
1742 ATAACTTACAGTGCAGGTCTGGATATGTATATGATACAGGCACTTCAGATCTGATTGTACA 1801
3106 GluLysGlyValTrpSerGlnProTrpValCysGluProLeuSerCysGlySerPro 3125
1802 GAGAAAGGGGTATGGAGCCAGCTTATCCAGTCTGTGAGCCCTTGTCTGTGGGTCCCA 1861
3126 ProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTrpGluSerGluValLys 3145
1862 CCGTCTCTGCCAATGAGTGGCAACTGGAGGGGACACACTATGAAGTGAAGTGAAG 1921
3146 LeuArgCysLeuGluGlyTrpThrMetAspThrAspThrAspThrPheThrCysGlnLys 3165
1922 CTCAGATGTCTGGAAGTATACCATGGATACAGATACAGATACATTCACCTGTCCAGAAA 1981
3166 AspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeuProGlu 3185
1982 GATGTCGCTGGTTCCCTGAGAGAAATCTCTGAGAGTCTCTAAATAATGTCTCTCCCGGAA 2041
3186 AsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerVal 3205
2042 AACATAACATATACTTGTGTCATGGGAGCAATTCAGTGTGAATAGCAAGTTCCTGTG 2101
3206 SerCysAlaGluGlyTrpThrPheGluGlyValAsnIleSerValCysGlnLeuAspGly 3225
2102 TCATGTGCAGAGGGTATACCTTTTGGAGGAGTTAAACATATCAGTATGTTCAGCTTCATGGA 2161
3226 ThrTrpGluProProPheSerAspGluSerCysSerProValSerCysGlyLysProGlu 3245
2162 ACCTGGAGGACCACTTCTCCGATGAATCTTGCAGTCCAGTTTCTTGTGGGAAACCTGAA 2221

3246 SerProGluHisGlyPheValValGlySerIysThrPheGluSerThrIleLeuTyr 3265
 |||||
 2222 AGTCCAGAACATGGATTGGTGGAGTAAATACACCTTTTGAAGCACAATTATTAT 2281
 |||||
 3266 GlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluValCysGlnGluAsnArg 3285
 |||||
 2282 CAGTGTGAGCTGGCTATGAATAGAGGGGGAACAGGGAACGTGTCTGCCAGGAGAACAGA 2341
 |||||
 3286 GlnTrpSerGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPhe 3305
 |||||
 2342 CAGTGGAGTGGAGGGTGGCAATATGCAAGAGACACCGAGGTGTGAACCTCCACTTGAATTT 2401
 |||||
 3306 LeuAsnGlyLysAlaAspIleGluAsnArgThrGlyProAsnValValTyrSerCys 3325
 |||||
 2402 CTCATGGGAAGCTGACATTGAAACAGAGCAGCTGAGCCCAACGTGGTATATTCCTGC 2461
 |||||
 3326 AsnArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrp 3345
 |||||
 2462 AACAGAGGCTACAGTCTTGGAGGGCCCTGAGGACACCTGCACAGAAATGGACCTGG 2521
 |||||
 3346 SerHisProValProLeuCysLysProAsnProCysProValProPheValIleProGlu 3365
 |||||
 2522 AGCCACCCAGTCCCTCTCGCAACCAATCCATGCCCTGTTCTTTGTGATTCGCCAG 2581
 |||||
 3366 AsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArg 3385
 |||||
 2582 AATGCTCTGCTGTCTGAAAGGAGTTTATGTGTGATCAGAAATGTCCCAATATGAG 2641
 |||||
 3386 GluGlyPheLeuLeuGlnGlyHisGlyIleThrCysAsnProAspGluThrTrpThr 3405
 |||||
 2642 GAAGTGTTCCTGTCAGGCGCCACGATCATCTCTGCACCCGACGAGCTGACGA 2701
 |||||
 3406 GlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAla 3425
 |||||
 2702 CAGACAAAGCCCAATGTGAAAAATCTCATGTGGTCCACAGCTCAGGTAGAAAAATGCA 2761
 |||||
 3426 IleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGly 3445
 |||||
 2762 ATTGCTCGAGGGGTACATTATCAATATGAGACATATCATCTCTGACCTCATCTGTTACGTGGA 2821
 |||||
 3446 TyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerPro 3465
 |||||
 2822 TACATGTTGGAGGGTTCTCTGAGGAGTGTGTTGTAGAAAATGACATGACATCACCT 2881
 |||||
 3466 ProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgPro 3485
 |||||
 2882 CCTATTTCAGAGCTGCTCTGTCGATTTCCATGTCAGATGGGGGCATCTGCCAACGCCCA 2941
 |||||
 3486 AsnAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGluGluProIleCysIle 3505
 |||||
 2942 AATGCTTGTCTGTCAGAGGCTGGATGGGGCGCTCTGTGTGAAGAACCATCTGCATT 3001
 |||||
 3506 LeuProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGly 3525
 |||||
 3002 CTTCCCTGTCTGAACGGAGTCTGTGTGGGCCCTTACCGTGTGACTGCCCGCTGGC 3061
 |||||
 3526 TrpThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLys 3545
 |||||
 3062 TGGAGGGGTCTGCTGTATACAGCTGTTGGCAGTCTCCCTGCTTAATGTGGAAAA 3121
 |||||
 3546 CysValArgProAsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArg 3565
 |||||
 3122 TGTGTGAAGACCAACCGATGCTACTGTCTTCTTCTTGGACGGGACATACTGTTCCAGG 3181
 |||||
 3566 LysArgArgThrGlyPhe 3571
 |||||
 3182 AAAAGGAGGACTGGGTGT 3199
 |||||

RESULT 22

AC096906

LOCUS

DEFINITION Rattus norvegicus clone CH230-11319, *** SEQUENCING IN PROGRESS

AC096906 266868 bp DNA linear HTG 10-MAY-2003

Rattus norvegicus clone CH230-11319, *** SEQUENCING IN PROGRESS

***, 3 unordered pieces.

AC096906

AC096906.6 GI:30522617

HTG; HTGS_Phrase1; HTGS_DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 266868)

REFERENCE

AUTHORS

Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Bayes, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoque, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. I., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X. Z., Sorrelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, S., Yen, J., Yoon, L., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

Direct Submission

Unpublished

2 (bases 1 to 266868)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (02-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 266868)

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:23267390.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGN7
Center clone name: CH230-11319
----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 223096 bases at least Q40
Consensus quality: 226986 bases at least Q30
Consensus quality: 229938 bases at least Q20
Estimated insert size: 232899; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 259563: contig of 259563 bp in length
* 259564 259663: gap of unknown length
* 259664 261157: contig of 1494 bp in length
* 261158 261257: gap of unknown length
* 261258 266868: contig of 5611 bp in length.

FEATURES

source
1..266868
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-11319"
misc_feature
1..1228
/note="wgs contig"
76973..78178
/note="wgs contig"
222699..222742
/note="clone_boundary"
clone_end:r7
site:ECORI
end_sequence:RH262152"

ORIGIN

Alignment Scores:
Pred. No.: 4,538-273 Length: 266868
Score: 5142.50 Matches: 1060
Percent Similarity: 40.61% Conservative: 127
Best Local Similarity: 36.26% Mismatches: 215
Query Match: 25.75% Indels: 1523
DB: 2 Gaps: 14

US-09-977-053-4 (1-3571) x AC096906 (1-266868)

Qy 1857 CyslleGluLeuAlaPheThrPheGlySerLysValThrTyrArgCysAsnLysGly 1876
||||| : : : : :
:::|

Db 113179 TGGCGAGAAACTTGGCTGGGTGGGGG-----CGTGGC 113214
Qy 1877 -----TyrThrLeuAlaGlyAspLysGlu 1884
Db 113215 CCGCTTTTGGGGCGTGGCTCGTATTTCTAGTACTGGGAAGTGGTGAACGAG 113274
Qy 1885 SerSerCysLeuAlaAsnSerSerTyrSerHisSerProValCysGluProValLys 1904
Db 113275 -----GTTAGCAGANTCTCGAGACTCACCAGTAAGCCAGCTGCT-ATGCTGCAA 113324
Qy 1905 CysSerSerProGluAsnLeuAsnAsnGlyLys----- 1915
Db 113325 TATTGAGTCCAGCAAAATAAAACAACAACAACAACAACAACAACAACAACA 113384
Qy 1916 -----TyrIleLeuSerGlyLeuThrTyrLeu 1924
Db 113385 TCAAAAGTAAGTAGTAGTACTCTTAATAATGATTGTGACCTCTGATCTCTCTCTCCA 113444
Qy 1925 SerThrAlaSerTyrSerCysAspThrGlyTyr-SerLeuGlnGly----- 1939
Db 113445 AGCATGTACATGTACACATGTACACTGCGGAGGAGCAAGCAAGCAACACACACAC 113504
Qy 1939 ----- 1939
Db 113505 ACACACACACACAGAGCAGACATACACACACACACACACACACACACACACAC 113564
Qy 1940 -----ProSer 1941
Db 113565 ACTCTCATTTGGAGTTCTTTAAATAGCAAAAGTCTTAAACCTATTTCCTTACTAATCGAGC 113624
Qy 1942 Ile---IleGluCysThrAlaSerGlyIleTyrAspArgAlaProProAlaCysHisLeu 1960
Db 113625 ATTAAATCATTTTACATAATAGTTCAATATTC-----AGTTGCTAAATA 113669
Qy 1961 ValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPheThr 1980
Db 113670 ATATTTTTTT-AAAAAACCCAAAGATAAAGGCAAGTAAGTGTGGAAGACTAGAAA 113728
Qy 1981 PheArgAsnThrValThrTyrThrCysLys----- 1990
Db 113729 ATGAAAGAGAGAGAGAAATGCTTTTGTGCGATCTACATGAGTTGTTTCTTCCCTTCCT 113788
Qy 1991 ---GluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrp 2009
Db 113789 TCTCATAGTACACCTTGGCGCCCTGACACCATCATATATGCCAGGCCAACGCACTGG 113848
Qy 2010 SerArgSerAspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHis 2029
Db 113849 AATTCAAGTAACCAACCAAGTGGCTGGCGCTCTCTCGCAGCAGGCCCCCAATGTGGACCAC 113908
Qy 2030 AlaSerProGluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAsp 2049
Db 113909 GCCTCTCCAGAGACTGGCAGACAGGCTCTTCGGGGACACCCCGTTCTACTACTGTGCGAGT 113968
Qy 2050 GlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValPro 2069
Db 113969 GGTTCAGCTGGCTGACAAATTCCTAGCTCATCTGCAACGCCAGGGGAACTGGGTTCCC 114028
Qy 2070 ProGluGlyGlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerVal 2089
Db 114029 CCGAGGGCCAGGCTGTGGCCCGCTGCATAGCTCACTTCTGTGAGAAACCCCACTCTGTT 114088
Qy 2090 SerTyrSerIleLeuGluSerValSerLysAlaIleAspPheAlaAlaGlySerValValSer 2109
Db 114089 TCTACAGCATCTTGGAGTCTGTGAGCAAGCAAAATTTTCAGCTGCTCGTGTGAGTANGC 114148
Qy 2110 PheLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGly 2129
Db 114149 TTCAAGTGCATGGAGGGTTTGTGACTGACACCTCAGGAGAGATTGAATGCTCGAGAGGC 114208
Qy 2130 GlyGlnTrpAsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluPro 2149
Db 114209 GGACAGTGGAGCCCTTCTCCCTTGTCCGTCAGTGCATCCAGTGGCGTGGGAGAGCCT 114268

QY	2150	ProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValIle	2169
Db	114269	CCAAAGCATCAAAATGGCTACCCAGTGGAACTACAGTTCGGGGCCGTGGGCG	114328
QY	2170	TyrSerCysAsnIleGlyPheThrIleGlyGlyIleValSerThrCysGlyAlaThr	2189
Db	114329	TACAGTCCCAAGAGATTCATATATCAAAAGGGGAGAAAGAGACATGTGAGGCCACG	114388
QY	2190	GlyGlnTrpSerProIleProThrCysHisProValSerCysGlyGluProProLys	2209
Db	114389	GGACAGTGGAGTAGACCCCTGCCACCTGCCACCTGTCCTGTAACGAACACCTTAAG	114448
QY	2210	ValGluAsnGlyPheLeuGlu	2216
Db	114449	GTTGGAACGGCTTCCTGGAGTAAGAGACTAGCAATGTCCATCCCTCATTTGGGTCACT	114508
QY	2216	-----	2216
Db	114509	ACCGTACTGAGTGAGTACTTCCAAAGTAAAGGACAGATTCGTAACAGAAAT	114568
QY	2216	-----	2216
Db	114569	ACTCTACACATATAGTCTACACATAGTGTGTGTAATAAATACTAATCAAGATAATTA	114628
QY	2216	-----	2216
Db	114629	GAACCTTACATATTTATCTTTTCAAAATCTTAGTTTATACATACAGATTTCCCTGC	114688
QY	2216	-----	2216
Db	114689	CCATATAACCATGCACCTGTATGTTCAAAGGAGTGAGAAAGGCATCAGATCCCATAGG	114748
QY	2216	-----	2216
Db	114749	ACTAGAGTTACACACAGTTGAGAGGCACCATGTGGGTGTAAAGATTCACACTTAGTCT	114808
QY	2216	-----	2216
Db	114809	CTGGAGGACAGCCGACTCTTAACCACTGTACCACTCTCCAGCCCCCTCAGTATACAT	114868
QY	2216	-----	2216
Db	114869	CTTTCTGCAATTATAGTAACCTATCGTTAGGCCTTTGTAACTTAGCTTAGCAATTAATG	114928
QY	2216	-----	2216
Db	114929	AAGTTAGAGCCGTTTTCATGAAATTCATTTGCTCTTACCAAACTCATCACCCTAAATG	114988
QY	2216	-----	2216
Db	114989	GAACCTCAGCGATGCTTATTGTTACTCCGATTTTAAAGGAACTGTGAGTGCAGTTTTTG	115048
QY	2216	-----	2216
Db	115049	CATTGCTCTTCGAGACATCTAAAGTAAATTTTACCACCGTGAGCCCGTTGTCAAGTGG	115108
QY	2216	-----	2216
Db	115109	GTACGAATGTAATCTGCTTCCCACTCCAGAGTGTCTTCTCACTAAGTGTGTGAGGAGNC	115168
QY	2216	-----	2216
Db	115169	TGGAACAAAAGCTGTAGTATAGTCTCTGGAGGTAAAGTCTCTCTGTGCAAGAGAGTGG	115228
QY	2216	-----	2216
Db	115229	TGAGCTTGATCTCCAGACCCAGGTAAGAAAAGCCAGGCAAGGTCACACCTATTA	115288
QY	2216	-----	2216
Db	115289	CCACAGCTCCCTGGGAGAGACACCTGACTCCCAAGTCTACTGCTGTGCTGGCCTAG	115348

Db	116429	::: TTGTAAGCTGGTGCCTTTGTCTCAGATCGCCCTTCCTTGATTCCCTTCGCGGTCGCCGGGTCTTC	116488
Qy	2395	rSerAlaleuHispheGlySerThrValIysenYrSerCysValGlyGlyPhePheLeuAr	2415
Db	116489	GGGTGCTCTCACITTTGGNAGTACCGTCAGATAITGTGTGTGTCGATGGTTTTCTTAAG	116548
Qy	2415	gGIYAasnSerThrrLeuCysGlnProAspGlyThrrTrpSerSerProLeuProGluCy	2435
Db	116549	AGGCAATCCAATCATCTCTGCCAAGTTGACGGCACCTGGAGTTCTCTCGTTGCCCGAGTG	116608
Qy	2435	sValProValGluCysProGlnProGluGluIleProasnGlyIlelleAaspValGlnGl	2455
Db	116609	CGTTCCGGTGGAAATGTCCTCCCAACTGAGGAGATCCTCAACGCGCATCATCTGTGCAAGG	116668
Qy	2455	yLeualatyrIleuSerThrrAlaLeuTyrrThrcysIysProGlyPheGluLeuValGlyAs	2475
Db	116669	GCTCGCCTTATCTCAGACAACCGCTCTACACCTGTAAAGCAGCGCTTCGAGTTAGTGGGCAA	116728
Qy	2475	nThrrThrrLeuCysGlyGluAsnGlyHisTripleuGlyGlyLysProThrcysLysAl	2495
Db	116729	CACCACCACTCTGTGGGGAATGGCCAGTGGCTTGGAGGAAAAACCAATGTGCAGACC	116788
Qy	2495	aileGluCysLeuIysProLysGluIleLeuasnGlyLysPheSerfyrThrrAspLeuHi	2515
Db	116789	CATTGAATGCCAGAGCCCAAGAAGATTCATAATGGCCNAATCTCTCCGTCGAGCTTTCA	116848
Qy	2515	sTrycIyGlnThrrValThrrYrSerCysAsnArgGlyPheArgLeuGluGlyProSerAl	2535
Db	116849	GTATGGACAAAACCATCATCACTCTTTGTGACGAGGCTTCOGGCTCGAAGGTCCCAATC	116908
Qy	2535	aLeuThrcysLeuGluThrrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHi	2555
Db	116909	CCTGACCTGTTTAGAGACAGTAACCTGGGATATGATGCCCATCTTGCAATGCCATCCA	116968
Qy	2555	sCysAaspSerProGlnProIleGluAsnGlyPheValGluGlyAlaAaspYrSerfyrGl	2575
Db	116969	CTGTAGTGACCCCCAGCCCATCGAAATGGTTTCGTAGAAGCGCGGATTCAGATATGG	117028
Qy	2575	yAlaIleIleIleTyrrSerCysPheProGlyPheGlnValAlaGlyHisalamerGlnTh	2595
Db	117029	CGCCATGANTATCTATATAGCTGTGTTCCTCCCGGTTTCAGGTGGTGGGCATGCCATGCAGAC	117088
Qy	2595	rCysGluGluSerGlyTrpSerSerSeriIeProThrcysMetProIleAaspCysGlyLe	2615
Db	117089	CTGCGAGAGAGCGGGTGGTCBAGCTCCAGCCGAGGTGTGTACCATAGACTGGGGTCT	117148
Qy	2615	uProProHisIleAaspPheGlyAaspCysThrrLysLeuLysAaspAsgGlnGlyTyrrPheGl	2635
Db	117149	CCCTCCTCACATAGACTTGGAGATCGCACTAGATCGAGCATGGCGCAGGATATTTTGT	117208
Qy	2635	uGlnGluAaspAspMetMetGluValProTyrrValThrrProHisProProTyrrHisLeuGl	2655
Db	117209	CCAGAAGACACATGATGGAAGTCCCATACTCTGACTCTCTCAC---CCTAACATTTGGA	117265
Qy	2655	yAlaValAlaLysThrrTpGluAsnThrrLysGluSerProAlaThrrHisSerSerAsnPh	2675
Db	117266	AGCGACGCCAAGCGTCGGAAATCACAGAGGAGTCCCTGTACCACATGCATCCCAATT	117325
Qy	2675	eLeuTyrgLyThrrMetValSerTyrrThrcysAsnProGlyTyrrGluLeuLeuGlyAsnPr	2695
Db	117326	TCTATATGCCACGCGGTGTCTATTCGCTGTGAGCTCTGTTATGAACTGCTGGGAATCCC	117385
Qy	2695	ovalleuIleCysGlnGluAaspGlyThrrTrpAsnGlySerAlaProSerCysyleIseRil	2715
Db	117386	TGTGCTGGTCTGCCAGGAAGATGGGCATCAATGAATGGGACCGCACCTCTTGTATTTCAT	117445
Qy	2715	eGluCysAaspLeuProThrrAlaProGluAsnGlyPheLeuArgPheThrrGluThrrSerMe	2735
Db	117446	TGAATGTGATTTGGCCGTGTGCTCCGAAAAATGGCTTTCTACATTTTACACAGACTACTAT	117505
Qy	2735	tGlySerAlaValGlnTyrrSerCysLysProGlyHisIleLeuAlaGlySerAaspLeuAr	2755

Qy	3115	oValCysGluProLeuSerCysGlySerProProSerValAlaAsnAlaValAlaThrGl	3135
Db	118643	AAcGTGTGAGCCCGTCTCTGTGTGGACCCACCACGGTGTCCAAATGCAGTGGCAACAGG	118702
Qy	3135	yGluAlaHisThrTyrGluSerGluValLysLeu	3146
Db	118703	AGAGGCACATACCTATTGAAAGCAAAAGTGAAACTCAGTAAAGACCATGGGTGGCTGTGGAG	118762
Qy	3146	-----	3146
Db	118763	TCGCAAGGAGTGTAGGTGAGAAGATTCTAAATGAAGTCAGAGGTCCAAACGCAGCTAA	118822
Qy	3146	-----	3146
Db	118823	TTAGAAAGAGAACCAACCATCTCTGGGCTGCCCTCCAAACCCCAAAAGCTTTTAAACATTTTA	118882
Qy	3146	-----	3146
Db	118883	TTTACTCGGACTATTGTAGAAATCCACTTACTATGAATAAATAATGTAGAACCTTTTCCAGA	118942
Qy	3146	-----	3146
Db	118943	CTACGGTTCTAAGCCCATCTCATGGTGCTCAAAAGCTGCAAGGTCTCTAAACCACCT	119002
Qy	3146	-----	3146
Db	119003	CAATTATGTTTGAATCACTCAGCAATAATGAAGAGGCTATCCACCCTGGCTACTTA	119062
Qy	3146	-----	3146
Db	119063	AGTCAGAATATGTTGCTCACATAATACTCCGCTGTCTGGGGGGATCGTGACATTAAT	119122
Qy	3146	-----	3146
Db	119123	ATTAATAATGAAGACGTAAAAATTTAGAATAAACTTAGGAATTAATAATGTTACTGCAAT	119182
Qy	3146	-----	3146
Db	119183	AAATAACTATCAGGTTCCAAAGCATCTCGATTGTCTCAGAATTCGTTCAATGTTCTGGG	119242
Qy	3146	-----	3146
Db	119243	AGAACTATGAGCAATTAATGCTCATACCTCATTCATTAATACTTAGCACCTCTCTGGGGTGG	119302
Qy	3146	-----	3146
Db	119303	CTCATTTGAAAGTTTTAGCCCCACCTTCAGGTTACCTGACTGTAGTAAAAAGCAAAGT	119362
Qy	3146	-----	3146
Db	119363	CTACGGAAGGACAGATGTTAGCCCAAAACAAAGAATGATGGGCCAGTCGGTGACTTTG	119422
Qy	3146	-----	3146
Db	119423	AAGTATATTTTAAGACTTCTGGGTAGATGTGATATGTTCAGGCACTTTGAGAGCCACGAT	119482
Qy	3146	-----	3146
Db	119483	GTCACCTGCTCATGAGGAGTCACTGGGGCTTTGCTTTGGCATGTGATGATGACCAATC	119542
Qy	3146	-----	3146
Db	119543	TTTGCTTAGTTTCCACTAATCAGGTTGGTTGGATTGTAGGAGACTTCTGTCTCCACCTA	119602
Qy	3146	-----	3146
Db	119603	ACCTGGTTAATAGTGGAAACTAAAGATATCTTCCAATGATCACTATAAGTTAATTA	119662
Qy	3146	-----	3146
Db	119663	TGGTTTAATTTGAAGGCAGCATACGGGAGTGTAGGAACATGGGGTCCAGCAATTCCT	119722

QY	3146	-----	3146
Db	119723	TCAGAGAACTCAGTCTCTCCACATCTCACATACCATGCGAGAGTCAATGGCTTCAAG	119782
QY	3146	-----	3146
Db	119783	AAGGGTGCTAACCTCATCGATGCCAAATGCGAGTGAGAGCCAAATGCGACCTGCTCATG	119842
QY	3146	-----	3146
Db	119843	AGGCAGAGCTGAGGATTGGACTCGTGGTATTGTGAAGCACACAGAGGGCTCACCAGAA	119902
QY	3146	-----	3146
Db	119903	AAGCATCCTTGTGTGGA CCCACTCACTGCATTAGTATTTACATTTTAAATCACTTGGGA	119962
QY	3146	-----	3146
Db	119963	ACTTAAGAGCCAAGTTGACTTAAGTGTAAACCATCTATGGCTATGACAGTGTAAACGA	120022
QY	3146	-----	3146
Db	120023	TCAAAGTTAAAGAGGACCTTATACAAGACCTCAGACAGTACTGAGTGTGAATAGGA	120082
QY	3146	-----	3146
Db	120083	GGCAAAACCATAAACACCATGTCTATCCCTATCAGTATCGGAGTCCGCTCCCTATAATTT	120142
QY	3146	-----	3146
Db	120143	CTTAAACTGGTGGCTATTTTCAGAAATGAACAGACAGCAAGTAACCCAGCAACATCTAAGACGA	120202
QY	3146	-----	3146
Db	120203	CTTTAGCCTTTGTGGTCTCGAGTGTGAGTCACTAAGAAATTTGTCTGGATGTGCTTTAT	120262
QY	3146	-----	3146
Db	120263	TAATGTCTTGTGAGGAGCGTTATCCATCTGACCTTTGAGCCAGTTTTCAGGTTTCTCAC	120322
QY	3146	-----	3146
Db	120323	AGAAATGCTGAGCCTGGGACAGTTGTGTTCACCTGGGCATAGGTGCGTCCACAGCGTGGT	120382
QY	3147	-----	3155
Db	120383	ACCTTCTTGGTTTTCATCTCTCTGTTTGGCATGCGAGTGTCTGGAAGGGTATGSGTGA	120442
QY	3155	pThrAspThrAspThrPheThrCysGlnLysAspGlyArgTrpPheProGluAArgIleSe	3175
Db	120443	TACGGACACAGACACATTCACCTGCCAGCAAGATGTCGGTGGTTCCTGAAAGNATCA	120502
QY	3175	rCysSerProLysLysCysProLeuProGluAsnIleThrHisIleLeuValHisGlyAs	3195
Db	120503	CTGCGAGTCTTAAACATGTCTGTGGCATCCACAGGACCGCATACGTGTTCCAGGAGA	120562
QY	3195	pAepPheSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrrThrPheGluG	3215
Db	120563	TGACTTCCAGGTGAATAGACAAAGTTCTGTGTGTGTCACAGAAAGGGTTTACCTATGACGG	120622
QY	3215	yValAsnIleSerValCysGlnLeu	3223
Db	120623	AGCAGACCGGTCAACGTGCGCAGGTAAAGTTCTGTCTGGCTTGGAATGTGTAAAGTTAACTC	120682
QY	3223	-----	3223
Db	120683	TGGCATCTGTTACTATGAACAGAACTCATGTCTTGGCCGCTGGAAGAAACAAATTCCTC	120742
QY	3223	-----	3223
Db	120743	AAGTTCTCACTTACATCAAGATTGAGATGCTTCCTTAGAAANTAGAAGTGGGGCAGGG	120802
QY	3223	-----	3223

Db 120803 GGATGGACTGGTGTATTAAGACACTTGTCTGCCCTCCAGAGACCTAGTGTTCGCCCCCA 120862
 QY 3223 ----- 3223
 Db 120863 GCACCCATATGTGGCTTACAGCTGTAACTACAGTGGGGGATCTGATGTCACCAAGCA 120922
 QY 3223 ----- 3223
 Db 120923 CAGAGAAATGCTGCGCAAAATACTACATGTATAGATAAAATTTGTTTAAACAGGAGGA 120982
 QY 3223 ----- 3223
 Db 120983 ARAAGAAATGAAGTGTGTTTTCGTAATATCCAGCTTTTAGATTAAAGAGGCC 121042
 QY 3223 ----- 3223
 Db 121043 AAATTTAAGGGGAATGTTAAACAGAAACAAACACTGGGCTGTAGAGATGCTCAGAG 121102
 QY 3223 ----- 3223
 Db 121103 ATTAAGAGCACTGCTGCTCTTCCAAAGGTCCTGAGTTCAATTTCCAGCAACTACATGGT 121162
 QY 3223 ----- 3223
 Db 121163 ACCTTACACCATGTGTAAATGGGATCCATGCCCTCTCTGCTGTGTCTGAAGACAGATA 121222
 QY 3223 ----- 3223
 Db 121223 CAGGTACTTTATAGAAATTAATTAACAGTTAATTAACAGTTAATTAATTAACCT 121282
 QY 3223 ----- 3223
 Db 121283 ATAGTGAGATATGATAGCAAGTATCCATGCCAGATGATCTGTATGAGAAACC 121342
 QY 3223 ----- 3223
 Db 121343 TGCCAGGGGCTGTGATGCTTAAGTAGAAGGCCAGATGTTGCAACAGCAGCAGCAA 121402
 QY 3223 ----- 3223
 Db 121403 CAGCAGGGTTCCTGGGATAAAAGTGGGGTATCTGCACCTAACCTCCCTCAGCCCAACC 121462
 QY 3223 ----- 3223
 Db 121463 TCTAGTAGCATGCTGTACTAAAGACTTCCTTATGGACACCTCTGTTTGGGGAATTCG 121522
 QY 3223 ----- 3223
 Db 121523 TTTATTTTGTCTCTGAGATGTCAACAGTGTTCAGTGTGCTGTTTAAAGACCT 121582
 QY 3223 ----- 3223
 Db 121583 AGATAGTGTAAAGAACCACTCAATTAACCCCTGGGTCCTTTTGCATCTCTCACTTTGGGA 121642
 QY 3223 ----- 3223
 Db 121643 TGAGTCTGCCATATAACACAGAACACCAACGATATAATCCCACTGACCTCTGCTGTCTG 121702
 QY 3224 ----- AspGlyThrTrpGluProPheSerAspGluSerCysSerProVal 3240
 Db 121703 CTCTTTAGCCGATGGTACATGGAGGCCACCGCTTCCGAGGAATCTGCACTCCGCTTG 121762
 QY 3240 erCysGlySerProGluSerProGluHisGlyPheValValGlySerLysThrPheG 3260
 Db 121763 TCTGTGGGAGCTCTGAAGTCCAGAGCATGGATTTCTGTGTGGCGGAGTGAATACAGCTTCG 121822
 QY 3260 luSerThrIleLeuGlnCysGluProGlyTyrGluLeuGluGlyAsnArgGluArg 3280
 Db 121823 GAGACCGGTGTTTATCACTGTGACCTCGCTGCTATGATTAAGAG-----G 121867
 QY 3280 alCys 3281
 |||||

Db 121868 TATGC 121872

RESULT 23
 LOCUS AL929406/c
 DEFINITION Mouse DNA sequence from clone RP23-332H16 on chromosome 4, complete sequence.
 ACCESSION AL929406
 VERSION AL929406.10 GI:31335532
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 152909)
 Pelan.S.
 Direct Submission
 Submitted (31-MAY-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Jun 2, 2003 this sequence version replaced gi:31076167.
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk

 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-332H16 is from the RPCI-23 Mouse BAC Library constructed by the group of Pieter de Jong.
 For further details see http://www.chori.org/bacpac/home.htm
 VECTOR: pBACe3.6.
 Location/Qualifiers
 1. 152909
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clones="RP23-332H16"
 /clone_lib="RPCI-23"

FEATURES
 source
 1. 152909
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="4"
 /clones="RP23-332H16"
 /clone_lib="RPCI-23"

ORIGIN
 Alignment Scores:
 Pred. No.: 2.78e-270 Length: 152909
 Score: 5086.50 Matches: 1008
 Percent Similarity: 45.60% Conservative: 117
 Best Local Similarity: 40.86% Mismatches: 183
 Query Match: 25.47% Indels: 1160
 DB: 10 Gaps: 13
 US-09-977-053-4 (1-3571) x AL929406 (1-152909)

QY	1993	TyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSer	2012	2216	-----	2216
Db	144647			143568	TTCTGCATTATAGTAGCCTATAGATAAGGCCTTTGTAGCCATATCTTAGCCATAAATGTG	143509
QY	2013	AspGlnGlnCysLeuAlaValSerCysAspGluProProIleValAspHisAlaSerPro	2032	2216	-----	2216
Db	144587			143508	AAGTTAGAGTGAGCCAATTTTGATGAAATTCATATTTTGTCTTTTATAAAGTTTCATGCGCTG	143449
QY	2033	GluThrAlaHisArgLeuPheGlyAspIleAlaPheTyrTyrCysSerAspGlyTyrSer	2052	2216	-----	2216
Db	144527			143448	AATGGAACTAAATAGATATTTATTTGTACTCTGATTTTAAAGTGAGATAGAGGTGCAGTT	143389
QY	2053	LeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrpValProProGluGly	2072	2216	-----	2216
Db	144467			143388	TTTGCAATTCCTCGTTCAAGATATCTAAAGTAATCTCACCATGGTGAGCCCAATTTGTTAAG	143329
QY	2073	GlnAspMetProArgCysIleAlaHisPheCysGluLysProProSerValSerTyrSer	2092	2216	-----	2216
Db	144407			143328	TGTGATACAAATGTGCTCTGCTCCACCCTAAAGTGTCTTCCCACTAAGTGTGTGCGGAAG	143269
QY	2093	IleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValValSerPheLysCys	2112	2216	-----	2216
Db	144347			143268	ACAAGAAACAGGGCTGTAGTGTCTCCAGACGTAAAGTGTCTTCTGTGCAAGAGCGAG	143209
QY	2113	MetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArgGlyGlnTrp	2132	2216	-----	2216
Db	144287			143208	ATGTGAGCTTGATCTCCAGAAATCCAGGGTAAGAAAGCCAGGCATGGTGCCACACACCTA	143149
QY	2133	AsnProSerProMetSerIleGlnCysIleProValArgCysGlyGluProProSerIle	2152	2216	-----	2216
Db	144227			143148	GAACCCAGGGCTGGCCAGGAGGACAGCCTGATCCCCAGAGCTCACTGCTTGCAGC	143089
QY	2153	MetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMetValAlaTyrSerCys	2172	2216	-----	2216
Db	144167			143088	CTTGCTGCTGATAGCAATAGTTAAAGACTGTCTCAAGAAACGAATATTCAAAACAA	143029
QY	2173	AsnLysGlyPheTyrIleLysGlyLysLysSerThrCysGluAlaThrGlyGlnTrp	2192	2216	-----	2216
Db	144107			143028	GGTGGGTGGCATCTGAGAAATAACACCTGAGGTTGACTTCTGCACTCCACACATGCAC	142969
QY	2193	SerSerProIleProThrCysHisProValSerCysGlyGluProProLysValGluAsn	2212	2216	-----	2216
Db	144047			142968	CCACATATAACAAAGAGAAAGGGTGAGTGAAGAGAGAGGAAACAGAGGGGGAGTA	142909
QY	2213	GlyPheLeuGlu	2216	2216	-----	2216
Db	143987	GGCTTCTCGA - GGTAAGAGGCTAGGAAAGATCCATACTCTTATAGGGCCA	143929	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143928	CCTACCATGATTGATTCGGTATTTTCCAGTAGAGGAGTGAACAGATTGATTGTTGAAC	143869	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143868	AGAAATACTCAAGTCTACACAGTGTACTGGATAAATAATGGATCAAGATAATTAAAGAA	143809	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143808	ACCTTACAATATACATCTTTCAAGTCTTATTTATGCGTATAGGTATGGCCTGCACA	143749	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143748	TATAGCATGTACCTGTGTGTGTCAGGGAGGCCAGAAAGGCGAGATGCCATAGGATC	143689	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143688	AGAGTTATAGACAGGTGTGAGGCACCATGTGTAGTGTGTGAGATTCAACCTTAGTCTCTG	143629	2216	-----	2216
QY	2216	-----	2216	2216	-----	2216
Db	143628	GAGGAGCAGCAGTGTCTTTAACCACTAGAGCCACCCCTCAAAATGCATCT	143569	2217	-----	2224

Db	142488	GTAGGCTGTA	CTGACTTTTCTTCTCCCTGTCATCCAGCACCACTGGCAGGACCTTTGAG	142429	Db	141408	GGTTTCGTAGA	AGGTGCGGATTACAGATACGGTGCCCATGATCATCTATAGCTGCTTCCT	141349
Qy	2225	SerGluValArg	TyrGlnCysAsnProGlyTyrIysSerValGlySerProValPheVal	2244	Qy	2585	GlyPheGln	ValAlaGlyHisAlaMetGlnThrCysGluGluSerGlyTyrSerSerSer	2604
Db	142428	AGCGAAGCA	AGGTTCAGTGCACCCAGGCTATAAGGCGACCGGAAAGTCTGTGTTTGT	142369	Db	141348	GGGTTTCAG	GTGTGTGTATGTCATGTCATGAGACCTGTGAGAGTGGGATGTCAGCTCC	141289
Qy	2245	CysGlnAla	AsnArgHisTyrHisSerGluSerProLeuMetCysValProLeuAspCys	2264	Qy	2605	IlleProThr	CysMetProIleAspCysGlyLeuProProHisIleAspPheGlyAspCys	2624
Db	142368	TGCCAAGCA	ATCGCCATCGCCAGCGACGCGCCCTCTGTCTGTCACCCCTCTCAACTGT	142309	Db	141288	AGGCCAACCT	GTGTACCATAGACTGGGTCTCCCTCTCATACATAGACTTTTGGTGACTGT	141229
Qy	2265	GlyIysPro	ProProIleGlnAsnGlyPheMetIysGlyGluAsnPheGluValGlySer	2284	Qy	2625	ThrIysLeu	IysAspAspGlnGlyTyrPheGluGlnGluAspAspMetMetGluValPro	2644
Db	142308	GGGAAACCC	CTCCCATTCAGATGGCTTTTGAAGAGGAGAACTTTGAAGTAGGGTCC	142249	Db	141228	ACTAAAGTC	AGAGATGGCCAGGGACATTTTGTATCAAGAAGATGACATGATGAAGTCCCA	141169
Qy	2285	LysValGln	PhePheCysAsnGluGlyTyrGluLeuValGlyAspSerSerTyrThrCys	2304	Qy	2645	TyrValThr	ProHisProProTyrHisLeuGlyAlaValAlaLysThrTyrGluAsnThr	2664
Db	142248	AAGGTCAG	TTTGTCTGAATAGAGGATATGACTCGTGTGGTGAATTTCTTGACTTGC	142189	Db	141168	TATCTGGCT	-----CACCCCTCAACATTTGGAAGCAACAGCTAGGCCCTTGGAAATACA	141115
Qy	2305	GlnIysSer	GlyIysTyrAsnLysLysSerAsnProIysCysMetProAlaLysCysPro	2324	Qy	2665	LysGluSer	ProAlaThrHisSerSerAsnPheLeuTyrGlyThrMetValSerTyrThr	2684
Db	142188	CAGAAATCT	GGCAATGAGTAAGACCAAGCCGAGGTGTGTCCCCACCACCAAGTGTGCA	142129	Db	141114	AAGGAGTGG	CTGCTTATGAACTGCTGGGAATCCCTGTCTGCTGATCTGCCAGGAAGTGGTAGC	141055
Qy	2325	GluProPro	LeuLeuGlnAsnGlnLeuValLeuIysGluLeuThrThrGluValGlyVal	2344	Qy	2685	CysAsnPro	GlyTyrGluLeuLeuGlyAsnProValLeuIleCysGlnGluAspGlyThr	2704
Db	142128	GAGCCTCT	CTCTTAGAAAACCAAGCTCGTATTGAAGGAATTAGCTTCCGAGGTAGGAGTG	142069	Db	141054	TGGAGCTGG	TATGAACTGCTGGGAATCCCTGTCTGCTGATCTGCCAGGAAGTGGTAGC	140995
Qy	2345	ValThrPhe	SerCysIysGluGlyHisValLeuGlnGlyProSerValLeuIysCysLeu	2364	Qy	2705	TyrAsnGly	SerAlaProSerCysIleSerIleGluCysAspLeuProThrAlaProGlu	2724
Db	142068	ATGACCAT	TTTCTTGAAGGGGCGATGCTTGCAGGCGCCCTCTGTCTGGAAGTGTCTG	142009	Db	140994	TGGATGGT	ACCAGCACCTCTTGCAATTTCCATTTGCAATGTGATTTGCCCTGTGCTCCCGAA	140935
Qy	2385	ProLeuIle	SerPheGlyValProIleProSerSerAlaLeuHisPheGlySerThrVal	2404	Qy	2725	AsnGlyPhe	LeuArgPheThrGluThrSerMetGlySerAlaValGlnTyrSerCysIys	2744
Db	141948	CCCTTGAT	TCTCCCTGCGGTCTTCCGCTCTTCCGCTCTTCAATTTGGCAGTACTGTC	141889	Db	140934	AATGGCTTT	TATACATTTTACACAGACGACTATGGCGAGTGTGCACAAATATAGCTGCAAG	140875
Qy	2405	LysTyrSer	CysValGlyGlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnPro	2424	Qy	2745	ProGlyHis	IleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSer	2764
Db	141888	AAGTATCT	GTGTGTGTCAGCGGGTTTTTCTTAAGAGGCGATCCACCATCTCTGCGAGGCT	141829	Db	140874	CCGGGGCACA	TTCTAGAAGGCTCCCACTTAAGACTCTGTCTGCAAGTAAGCAGTGGAGT	140815
Qy	2425	AspGlyThr	TrpSerProLeuProGluCysValProValGluCysProGlnProGlu	2444	Qy	2765	GlyAlaSer	ProArgCysGluAlaIleSerCysIysLysProAsnProValMetAsnGly	2784
Db	141828	GATAGCAC	CTGGAGTTCTCCATTGCCCGAATGCTTCCGCTAGAAATGTCCCAACCTGAG	141769	Db	140814	GGCACTGTT	CCACGCTGTGAAGCCATCTCATGCAAGTAAGCCAAACCCACTCTGGAATGGA	140755
Qy	2445	GluIlePro	AsnGlyIleIleAspValGlnGlyLeuAlaTyrIleuSerThrAlaLeuTyr	2464	Qy	2785	SerIleLys	GlySerAsnTyrThrLysSerThrLeuSerThrLeuTyrGluCysAspProGly	2804
Db	141768	GAGATCCT	CAACGGTATCATCCACGTACAAAGGCTTGCCTATCTCAGCACCAACGCTCTAC	141709	Db	140754	TCCATCAAG	SAGATGACTACTCTCTACCTGGGTGTGTATATACAGAGTGTGACTCTGCG	140695
Qy	2465	ThrCysLys	ProGlyPheGluLeuValGlyAsnThrThrThrLeuCysGlyGluAsnGly	2484	Qy	2805	TyrValLeu	AsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrpAspGluAsp	2824
Db	141708	ACCTGAAG	CCAGCGCTTTGAGTTAGTGGGCAATGCTACCAACCTCTGTGTGGGAAATGGC	141649	Db	140694	TATATCTCA	ATGGCTCTAAGAAGAGGACATGCCAAGAAATAGAGATTGGGATGGGCAT	140635
Qy	2485	HisTyrLeu	GlyGlyLysProThrCysLysAlaIleGluCysLeuLysProLysGluIle	2504	Qy	2825	GluProIle	CysIleProValAspCysSerSerProProValSerAlaAsnGlyVal	2844
Db	141648	CAGTGGCT	CGAGGAAACCAATGTGCACACCAATTTGAATGCCAGAGCCCAAGAGATT	141589	Db	140634	GAGCCCATG	TGTATTCTGTAGACTGTGGCTCACCCCGAGTCCCCACCAATGGCCGAGTG	140575
Qy	2505	LeuAsnGly	IysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCys	2524	Qy	2845	ArgGlyAsp	GluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGlyPheLeu	2864
Db	141588	TTAAATGG	CCAAATCTCTCCGCTGAGCTTTTCAGTATGACACCAACCATCATCTTTTGT	141529	Db	140574	AAGGAGAGA	ATACATTTCCAAAAGAGATATACATACTCTTGCCTGGAAGGGTTTCATA	140515
Qy	2525	AsnArgGly	PheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrp	2544	Qy	2865	LeuGluCl	ValArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyValThrPro	2884
Db	141528	GACCGGGCT	CTCCGGCTCGAAGGTCCCAAAATCCCTGACCTCTTTAGAGACAGGTGACTGG	141469	Db	140514	CTGGAAGAG	CCAGGAGTGTATCTGTCTTACCAATGGAGTGGAGTGGGAGTGGCACTCCC	140455
Qy	2545	AspValAsp	AlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsn	2564	Qy	2885	AspCysVal	ProValArgCysAlaThrProProGlnLeuAlaAsnGlyValThrGluGly	2904
Db	141468	GATATGAT	CCCCCTCTTGTGATGCCATCCACTGCACTGACGTACCCACAGCCCATTTGAAAT	141409	Db	140454	AGCTGCAT	GTGTGTAGATGTCTGCGCCACCAACAGGTGCCAATGGGTGGCAGATGGC	140395
Qy	2565	GlyPheVal	GluGlyAlaAspTyrSerTyrGlyAlaIleIleIleTyrSerCysPhePro	2584	Qy	2905	LeuAspTyr	GlyPheMetLysGluValThrPheHisCysHisGluGlyTyrIleLeuHis	2924
Db					Db	140394	CTAGACTAT	GGGTTCAGAAGGAAGTAGCGTCTCCACTGTCTAGAGGGGTATGTGTGTCAG	140335
Qy					Qy	2925	GlyVala	ProLysLeuThrCysGlnSerAspGlyAsnTrpAspAlaGluIleProLeuCys	2944
Db					Db	140334	GGGGCTCCA	AGACTACCTGTCTAGTCCCAATGGGACTTTGGGATGGCAGAGTCCCTGTCTGT	140275

QY	2945	LysProValAsnCysGlyProProGluAspLeuAlaHisGlyPheProbenGlyPheSer	2964	3145	-----	3145
Db	140274	AAACCAAGCTACCTGTGTCTCTCCGACCTTCCCGAGGGCTCCCTAAATGGCTTTTCT	140215	139197	CTCAGTTGTCTCAAGATCTCGTTTACGTTCCCGGAGAACTACGGSCATTTATCAGCTCTTT	139138
QY	2965	PheIleHisGlyGlyHisIleGlnCysPheProGlyTyrIleLeuHisGlyAsn	2984	3145	-----	3145
Db	140214	TTTTATCATGGGGGCCACATACATCATCATGTGTGTCTTACTGGTATATAGCTTCATGGAAC	140155	139137	ATCTGATTGTAACTTAGCACCCTCACCTGGGGTGATCCATTCTGAACGTTTTTAGCCCCAC	139078
QY	2985	SerSerArgArgCysLeuSerAsnGlySerTrpSerGlySerProSerCysLeuPro	3004	3145	-----	3145
Db	140154	CCATCAAGAGATGCTTCCCAATGGTTCTCGAGGGCAGCTCGCCATCTCGCTACCT	140095	139077	CTTCAATGTTGCCCTGACTGTAGTAAAGCAAAAGTTTATGGAATATGGAATGTTAGTG	139018
QY	3005	CysArgCysSerThrProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCys	3024	3145	-----	3145
Db	140094	TGCAGGTGTTCACACCCATCATTCACAGGGAACCATCAACGCACTGANTTTGGATGT	140035	139017	GCTGAACAGTACCAGAGACACAGCACAAACCAAAAGGGTGAATAGCCAGTCTGT	138958
QY	3025	GlyLysAlaIleArgIleGlnCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIle	3044	3145	-----	3145
Db	140034	GGAAAGACGGTCCAGATTGAGTCTCAAGGCTTCAAGCTCTTGAGCTTTCTGAAATC	139975	138957	GCCTTTGGAGTAAATTTTAAACACTTCCGTTGCGATGTTGATGTCAGGCATTCTGAGAG	138898
QY	3045	ThrCysGluAlaAspGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCys	3064	3145	-----	3145
Db	139974	ACCTGTGATGCCAATGGCAATGGTCTGAC---GTCCCACTGTGTGAGCACTCAGTGC	139918	138837	CCACAATGTCACTCTCTCGTCAGGACTCACCTGGGGCTTCATCTTTTGGCATGTGATGATG	138838
QY	3065	GlySerLeuProMetIleProbenAlaPheIleSerGluThrSerSerTrpLysGluAsn	3084	3145	-----	3145
Db	139917	GGCGCTCTCCCAACCATACCCCAACCAATTTGCTTGAGGGCAGGCTTTTCGGAGGACAA	139858	138837	ACCAATCTTTGCTGTGGTTTCCGCTGATCAGGTTGTTCTGATTTGTACGAGACTCTCTCTC	138778
QY	3085	ValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCys	3104	3145	-----	3145
Db	139857	GTGGTAACCTTACGCTGCGAGCTGGCTACACCATGCAAGGTAGTTCAGATCTGATTGT	139798	138777	CCACCTAGCTGGTTAATATATGTGAATTTAAAGGTGCTCTCCAATGATCACTATGAAAT	138718
QY	3105	ThrGluLysGlyValTrpSerGlnProTyrProValCysGluProLeuSerCysGlySer	3124	3145	-----	3145
Db	139797	ACGGAAAAGCGATATAGAGCCAGCTTACCCAACGTGTGAACCCCTGCTCTGTGGACCC	139738	138717	AACTAAATGTCTGATATGAAGCCAGCCTATGGGAGTGGTAGAAACATGGGTACCAGAG	138658
QY	3125	ProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluVal	3144	3145	-----	3145
Db	139737	CCACCAACTGTAGCCAATGTCAGTGGCAACAGGAGGCTCATACCTATGAAAGCAAGTG	139678	138657	ACTCCCTCTAGAGAACACAGTTTCTCCACATCTCACACTACCACCGCAGAGTCATGTGGC	138598
QY	3145	Lyb-----	3145	3145	-----	3145
Db	139677	AACTCAGTAAGACTATGGTGGCTGTGGAACTTCAAGGAGTGTGAAGCAAGATTA	139618	138597	TTCAAGAGGGGTGCCAACCTCATCAATGCCAAATGGCAACAGGGCCCTAAATGGCACCTG	138538
QY	3145	-----	3145	3145	-----	3145
Db	139617	TAAATGAAGTCAGAGGTGAAATGTAGCTAATTAGAAGATAACCGTCATCCCTGGGCT	139558	138537	CTCATGAGGTAGGCATGAGAAATTTGATGAGTTGATGTTTGTAAAGCACACAGGGGAGTCA	138478
QY	3145	-----	3145	3145	-----	3145
Db	139557	GCTCTCCAGACCTCAAGCTTTTAAACATTTATTTACTCAGGACTATGTACAAATCCATT	139498	138477	CCAGAAATAGTTATTGTATGGAACCCAGTTGCTGTCATAGTATTTAAATTTGTAAATCAC	138418
QY	3145	-----	3145	3145	-----	3145
Db	139497	TACTATGAAATTAATGTATACCTTTCCAGGCTGTGGTTCTTAATCAATCCCTCATGGTGC	139438	138417	ATGGAGACTTAAGAGCCAAATTTGACCTAAAGCTTTAACTACTCTAGTACTATGACAGTGT	138358
QY	3145	-----	3145	3145	-----	3145
Db	139437	TCACAAAGCGCAAGTCTCTTAACCGAGCTCAATTTATGTGTGAATCACTGAGCAATATA	139378	138357	TAATGACCAAGTTTAAAGGAGTACCTTGTGCAAGCACTCAGACAGTACTTGTGTCACA	138298
QY	3145	-----	3145	3145	-----	3145
Db	139377	AATAAGGATGCTATCCAGCTGTGATATTAAAGCTAAGAAATATGGTGTGCATATAATGA	139318	138297	GTAGACAGCCATAAACACTATGTCTATTCCTATCATGTATCTGAGTCCATTTTGAACCT	138238
QY	3145	-----	3145	3145	-----	3145
Db	139317	AACTCCCTGTCTGAAGGAATTATACATTAAATTTAAATGAAGACATGAAAAATT	139258	138237	CTCTATAATTTCTTGAAGTACTAGTACTATTTTCAAGAAATGGCAAAATGATAAGGGAAGTAAT	138178
QY	3145	-----	3145	3145	-----	3145
Db	139257	TGGAATGAACCTTAGGAATTTAAATTTTATTTGGAATTAATAGACACCATGTTAAAGGCAT	139198	138177	TTTCATCTAAACACCTTAGCTAGCTAAGAAATTTATCTGGCTGTCTGTTTAAATGTTTGTG	138118
QY	3145	-----	3145	3145	-----	3145

Db 138117 CAGTGATGTTATTTCATCTGACCTTTGAGCGGCTCTCAGGTTTCTCATAAAATGCTCTCCT 138058
 Qy 3145 ----- 3145
 Db 138057 GAGCCTTGTGGCAGTTCTGTTCATCTGAGCAGTGGCCAGTTACAGTGTGTACTATTG 137998
 Qy 3146 -----Leu-ArgCysLeuGluGlyTyrThrMetAspThrAs 3157
 Db 137997 TAGGTGTTCATTCTTCTGTTGTCATTGTCAGGTGTCTGGAGGTATGTGATGATTCGGA 137938
 Qy 3157 pThrAspThrPheThrCysGlnLysAspGlyArgTyrPheProGluArgLysSerCysE 3177
 Db 137937 TACAGATACATTACCTGCCAGCAGATGCGCATTTGGTCTCTGAAAGATCACCTCGAG 137878
 Qy 3177 rProLysLysCysProLeuProGluAnlleThrHisIleLeuValHisGlyAspAspPh 3197
 Db 137877 TCCTAAAAATGCCCTGTGCTGCCATCCACATGACACGATAGCTTTTCCCGAGATGACTT 137818
 Qy 3197 eSerValAsnArgGlnValSerValSerCysAlaGluGlyTyrThrPheGluGlyValAs 3217
 Db 137817 CAGGTGACACGACAGTTCTGTGTCATGTGCGAAGGGTTTACCCACGAGAGTGAA 137758
 Qy 3217 nIleSerValCysGlnLeuAspGlyThrTrpGluProProPheSerAspGluSerCysE 3237
 Db 137757 CTGGTCAACATGCCAGGTGAGTACTGTGTCGACCT-----GAGACGTGTAA 137710
 Qy 3237 rProValSer----- 3240
 Db 137709 GCCTAACTCTGGCATCTGTTACTATGACCAAACTCACATCTTGCTCTCTTGGAAAGAAC 137650
 Qy 3240 ----- 3240
 Db 137649 AATTCCTCCAGTTTTCATTACATCCAAAGATTAGAGTGTTCCTAGAAATAGAGGG 137590
 Qy 3241 -----CysGlyLysProGluSerProGluHisG 3250
 Db 137589 TGGGGCTGGGATGGTGGTTCAGAGCATGTGCTGGCCCT-----CCAGAGGACCT 137536
 Qy 3250 yPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluProG 3270
 Db 137535 AGGCTTGGTTCCTCCAGCACCCAC-----ATGGTAGCTTACAACTGC-----AA 137494
 Qy 3270 yTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSerGlyV 3290
 Db 137493 CTACAGG---TCAGGGACCGGATGCTCATATGCACAGAGAAAGCT-----GCAGGCAA 137443
 Qy 3290 yValAlaIleCysLysGluThrArgCysGluThr----- 3301
 Db 137442 AATACTCACATGTATAAAATAAAATTTGTTTAAACAGGCGAGAAAGAAATAGAACTTGG 137383
 Qy 3302 -----ProLeu-----GluPheLeuAsnGlyLysAlaAspI 3312
 Db 137382 GTATTTTTCCTAATATTCCTCAGTGTAGATTAGAGGAGTTAACTTAAGGGAATGT 137323
 Qy 3312 eGluAsnArgThrThrGly 3318
 Db 137322 TAAAGCAAAACACAGGG 137304

RESULT 24
 AX540417
 LOCUS AX540417 3262 bp DNA linear PAT 23-NOV-2002
 DEFINITION Sequence 29 from Patent WO02055738.
 ACCESSION AX540417
 VERSION AX540417.1 GI:25273450
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Panzer,S.R., Lincoln,S.E., Altus,C.M., Dufour,G.B., Hillman,J.L.,
 Jones,A.L., Dam,T.C., Liu,T.P., Harris,B., Flores,V., Daffo,A.,

TITLE Marwaha,R., Chen,A.J., Chang,S.C., Gerstin,E.H., Peralta,C.H.,
 JOURNAL David,M.H. and Lewis,S.A.
 FEATURES Molecules for disease detection and treatment
 source Patent: WO 02055738-A 29 18-JUL-2002;
 1. 3262 Incyte Genomics, Inc. (US)
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: LI:202943.4:2001JAN12"

ORIGIN

Alignment Scores: Pred. No.: 8.68e-264 Length: 3262
 Score: 4930.00 Matches: 947
 Percent Similarity: 95.33% Conservative: 12
 Best Local Similarity: 94.14% Mismatches: 20
 Query Match: 24.68% Indels: 28
 DB: 6 Gaps: 4
 US-09-977-053-4 (1-3571) x AX540417 (1-3262)
 Qy 1590 LysSerLeuAlaThrSerCysProGluGluLeuSerLysGlyAsnValLeuAlaTrpPro 1609
 Db 289 GAGTCACTGGCTACTCTCTGCCAGAGGAACCTCAGTAAAGGAAACGCTGTAGCATGGCT 348
 Qy 1610 AspPheLeuSerGlyIleValGlyValValValValValValValValValValValVal 1629
 Db 349 GATTTCCTGTGAGGAATTTGGGGGAAAGTGAAGATCGATTCTAAGAGCATATTTTGTCT 408
 Qy 1630 AspCysProArgLeuGlyGlySerValProHisLeuArgThrAlaSerGluAspLeuLys 1649
 Db 409 GGTGGCCACGCTTAGGGGGTCACTGAGAGCTCATCTGAGAGCTGATCTGAAGATTAAAG 468
 Qy 1650 ProGlySerLysValAsnLeuPheCysAspProGlyPheGlnLeuValGlyAsnProVal 1669
 Db 469 CCAGGTCCAAAGTCATCTCTTCTGTGATCCAGGCTTCCAGCTGGTGGGAAACCTGTG 528
 Qy 1670 GlnTyrCysLeuAsnGlnGlyGlnTrpThrGlnProLeuProHisCysGluArgLysSer 1689
 Db 529 CAGTACTGTCTGAATCAAGACAGTGGACACACACCTTCTCTACTGTGACGATTAAG 588
 Qy 1690 CysGlyValProProProLeuGluAsnGlyPheHisSerAlaAspAspPheTyrAlaGly 1709
 Db 589 TGTGGGTGGCCACTCTCTTGGAGATGGCTTCCATTCCAGCGGATGACTTCTATCTGGC 648
 Qy 1710 SerThrValThrTyrGlnCysAsnAsnGlyTyrTyrLeuLeuGlyAspSerArgMetPhe 1729
 Db 649 AGCACATGAATACACAGTGCACATGGCTTACTATCTATTGGGTGACTCAGGGATGTC 708
 Qy 1730 CysThrAspAsnGlySer-TrpAsnGlyValSerProSerCysLeuAspValAspGluC 1749
 Db 709 TGACATGATATGGAGCTAGGAACGGGTTCACCATCTCCCGCTGATGTCGATGAGTG 768
 Qy 1749 sAlaValGlySerAspCysSerGluHisAlaSerCysLeu-AsnValAspGlySerTyrI 1769
 Db 769 TGCAGTTGGATCAGATTGTAGTGAGCATGCTTCTTGCTGCTGAACGCTAGATGATCTACA 828
 Qy 1769 leCysSerCysValProProTyrThrGlyAspGlyLysAsnCysAlaGluProLysC 1789
 Db 829 TATGTTTCATGTGTCCTCCCGCTACACAGAGATGGGAAAGAACTGTCAGAACCTATAAAT 888
 Qy 1789 ys-LysAlaProGly---AsnProGluAsnGlyHisSerSerGlyGluIleTyrThr-Va 1807
 Db 889 GCTAAGGCTCCAGCGCAGAAATCCGGAATATGCGCACCTCTCAGGTGAGATTATACAGT 948
 Qy 1807 lGlyAla---AlaValThrPheSerCysGlnGlyTyr-GlnLeuMetGlyValThrL 1826
 Db 949 AGGTGCCCGGAGTGCACATTATTACGTGTGAGGAAGGATACCAGTGTGAGGAGTAACA 1008
 Qy 1826 ysIleThrCysLeuGlu-SerGlyGluTrpAsnHisLeuIlePro-TyrCysLysAlaVa 1845

Db 1009 AAATCACATGTTGGAGTACTGGAGAAATGGAATCATCTAATACCAATATTTGTAAGCTGT 1068
Qy 1845 lSerCysGlyLysPro-AlaIleProGluAsnGlyCysIleGlu-GluLeuAla---Phe 1863
Db 1069 TTCAATGGTAAACCGGACTATTCCAGAAATATGTTGCATTGACGGAGTTAGCCACTTTT 1128
Qy 1864 ThrPhe---GlySerLysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAsp 1882
Db 1129 ACCTATTTGGGCGCAAGTGCATATAGTGTATAAATAAGGATATATCTCTGGCGGTGAT 1188
Qy 1883 LysGluSerSerCysLeuAlaAsnSerTrpSerHisSerProProValCysGluPro 1902
Db 1189 AAGAAATCACTCTGCTTACAGTCTCTGGAGTCATTTCCCTCTCTGTGTGTGAACCA 1248
Qy 1903 ValLysCysSerSerProGluAsnIleAsnAsnGlyLys-TyrIleLeuSerGlyLeuTh 1922
Db 1249 GTGAAGTGTCTAGTCCGGGACAATAATACTAATGGACAACCTATATATAGATGGGCTTAC 1308
Qy 1922 rTyrLeuSerThrAlaSer-TyrSerCysAspThrGlyTyrSerLeuGlnGlyProSer- 1941
Db 1309 CTACCTTTCTACTGCATCATATAATTATGCGATACAGGATACAGCTTTACAGGGCCCTTTCCC 1368
Qy 1942 IleIleGluCysThrAlaSerGlyIleTrpAspArgAlaPro-ProAlaCysHisLeuVa 1961
Db 1369 ATATTGGAATGACGGCTTCTGGCATCTGGACAGAGCGGACCCCTGCTGTGCACCTCGT 1428
Qy 1961 lPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyAsnAsnPhetrPh 1981
Db 1429 CTTCTGTGGAGAACCACTGCCATCAAGATGCTGTCAATTACGGGGAATAACTTCACTTT 1488
Qy 1981 eArgAsnThrValThrTyrThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIl 2001
Db 1489 CAGGAACACCGTCACTTACACTTGCACAAAGAGGCTATACTCTTGTGTGTCTTGACACCA 1548
Qy 2001 eGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCy 2021
Db 1549 TGAATGCTTGGCGGACGGCAGTGGAGTAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1608
Qy 2021 sAspGluProProIleValAspHisAlaSerProGluThrAlaHisArgLeuPheGlyAs 2041
Db 1609 TGATGAGCCACCATTTGTGGACCAAGCCCTCTCCAGAGACTGGCCCATCGGCTCTTGGAGA 1668
Qy 2041 pileAlaPheTyrTyrCysSerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCy 2061
Db 1669 CATTCGATTTCTACTGCTCTGATGGTTACAGCTTACAGCTTACAGCAATTCCTCCAGCT 1728
Qy 2061 sAsnAlaGlnGlyLysTrp-ValProProGluGlyGlnAspMetProArgCysIleAlaH 2081
Db 1729 CAATGCCCAAGGCGAGTGGGTACCCCAAGAGGTCAAGACATGCCCCCTTGTATAGCTC 1788
Qy 2081 iAPheCysGlu-LysProProSerValSerTyrSerIleLeuGluSerValSerLysAla 2100
Db 1789 ATTCTGTGAAAACCATCTCATCGTTCCTATAGCATCTTGGAAATCTGTGAGCAAAAGCA 1848
Qy 2101 LysPheAlaAlaGlySerValValSerPheLysCysMetGluGlyPheValLeuAsnThr 2120
Db 1849 AAATTTGAGCTGGCTCAGTTGTAGCTTTAAATGATGATGAGAGGCTTTGTACTGAAACACC 1908
Qy 2121 SerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMetSerIleGln 2140
Db 1909 TCAGCAAAAGATTGAATGATGAGAGTGGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1968
Qy 2141 CysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyrAlaSerGlySer 2160
Db 1969 TGCATCCCTGTGCGGTGTGGAGAGGCCAACCAAGCATCATGAAATGGCTTATGCAAGTGGATCA 2028
Qy 2161 AsnTyrSerPheGlyAlaMetValAlaTyrSerCysAsnLysGlyPheTyrIleLysGly 2180
Db 2029 AACTACAGTTTGGAGCCATGSGGTCTTACAGCTGCAACAGGGGTCTTACATCAAAAGGG 2088
Qy 2181 GluLysLysSerThrCysGluAlaThrGlyGlnTrpSerSerProIleProThrCysHis 2200
Db 2089 GAAGAAGAGCACTCTGCCAGGCCACAGAGGAGTGGAGTGTAGTCTTATACCGAGCTGCCAC 2148

Qy 2201 ProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHisThrThrGly 2220
Db 2149 CCGGTATCTTGTGTGTAACACCACTAAGTTGAGAATGGCTTTCTGGAGCATACACTGGC 2208
Qy 2221 ArgIlePheGluSerGluVal-ArgTyrGlnCysAsnProGlyTyrLysSerValGlySe 2240
Db 2209 AGGNTCTTTGAGAGTGAAGTTGAGGTATCAGTGTATAACCCGGGTATAAGTCACATCGGAAG 2268
Qy 2240 rProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeuMetCysVa 2260
Db 2269 TCTGTATTGTCTGCCAAGCCCAATGCCCATCTGGCACAGTGAATCCCTCTGATGTGTGT 2328
Qy 2260 lProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGlyGluAsnPh 2280
Db 2329 TCTCTCGACTGTGAAAACCTCCCGCATCCAGATGGCTTCATGAAGAGAGAAAACCTT 2388
Qy 2280 eGluValGlySerLysValGlnPhePheCysAsnGluGlyTyrGluLeuValGlyAspSe 2300
Db 2389 TGAAGTAGGGTCCAGGTTTCAGTTTTCTGTGAATGAGGGTTATGAGCTTGTGTGTGACAG 2448
Qy 2300 rSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLysCysMetPr 2320
Db 2449 TCTTGGACATCTCAGAAATCTGGCAATGGAATGAAGTCAAAATCCAAAGTGCATGCC 2508
Qy 2320 oAlaLysCysProGluProProLeuLeuGlnLeuValLeuLysGluLeuThrTh 2340
Db 2509 TGCCAAGTCCAGAGCCGCCCTCTTGGAAAACCAAGCTAGTATTAAAGAGGTGGACCAC 2568
Qy 2340 rGluValGlyValValThrPheSerCysLysGlyGlyHisValLeuGlnGlyProSerVa 2360
Db 2569 CAGGTAGAGGTGTGACATTTTCTGTAAAGAGGGCATGTCTGCAAGGCCCTCTGT 2628
Qy 2360 lLeuLysCysLeuProSerGlnTrpAsnAspSerPheProValCysLysIleValLe 2380
Db 2629 CCTGAAATGCTTGGCATCCAGCAATGGAATGACTCTTTCCCTGTTGTAAGATGTTCT 2688
Qy 2380 uCysThrProProProLeuIleSerPheGlyValProIleProSerSerLalaLeuHisPh 2400
Db 2689 TTGTATCCCACTCCCTTAATTTCTTGTGTGTGCCCATTTCTCTCTCTCTCTCTCTCT 2748
Qy 2400 eGlySerThrValLys-TyrSer-CysValGlyGlyPhePheLeu-ArgGlyAsnSerTh 2419
Db 2749 TGAAGTACTGTCAAGGTATTTCTGATGTAGTGGGTGTTTTCTTACAGCAGGAATTTCTAC 2808
Qy 2419 rThrLeuCysGlnProAspGlyThrTrp-SerSerProLeuProGluCysValProValG 2439
Db 2809 CACCTCTGCCAACCTGTATGGCACCTGGAGGCTCTCCACTGACAGAAATGTTCCAGTAG 2868
Qy 2439 luCysProGlnProGluGluIleProAsnGlyIleIleAspValGlnGlyLeuAlaTyrIL 2459
Db 2869 AATGTCCCCCACTTGAGGAATCCCCCAATGGAATCATTTGATGTGCAAGGCTTTGCCCTATC 2928
Qy 2459 euSerThrAlaLeuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrIL 2479
Db 2929 TCAGCACACTCTCTATATCTGCAAGCCAGGCTTTGAAATTTGGTGGGAATTTACTACACC 2988
Qy 2479 euCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCys- 2498
Db 2989 TTTGTGGAGAAAATGGTCACTGGCTTGGAGGAAAACCAAGATGTAAGGCCATTTGAGTGCC 3048
Qy 2499 LeuLysProLysGluIleLeuAsnGlyLysPheSerTyrThrAspLeuHisTyrGlyGln 2518
Db 3049 GTGAAACCCCAAGAGATTTTGAATGGCAAAATCTCTTTACACGAGCCCTTACATATGGACAG 3108
Qy 2519 ThrValThrTyrSerCysAsn-ArgGlyPheArgLeuGluGlyProSerAlaLeuThrCy 2538
Db 3109 ACCGTACTACTCTTGTGCAACAGAGGCTTTGGCTCGAAGGTCCCAAGTCCCTTGACCTG 3168
Qy 2538 sLeuGluThrGlyAspTrpAspValAsp-AlaProSerCysAsnAlaIleHisCysAspS 2558
Db 3169 TTTAGAGACAGGTGATTTGGATGTAGATTGCCCATCTTTCGAATAGCATCCACTGTGATT 3228


```
QY 421 LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
DB 1594 CTACCCAAATGTTGTGGTTCGGCTTCAGAGAGCTACTGCAGAGTAAGAATGCTCTCAT 1653
QY 441 LeuArgGlnProLysHisGlyHisLeuSerCysSerThrArgGluMetLeuTyrLysThr 460
DB 1654 CTCGGCCAGCCGGAACATGCGCCACATCAGCTGTTCTACAGGGAATGTTATATAGACA 1713
QY 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
DB 1714 ACATGTTTGGTTCCTGTGATGAAGGTCACAGACTAGAGGCGAGTGATAGCTTACTTGT 1773
QY 481 GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr 500
DB 1774 CAAGGAAACACGAGTGGGATGGGCCAGAACCCCGGTGTGTGAGCGGCCACTGTTCACC 1833
QY 501 PheGlnMetProLysAspValIleLeuSerProHisAsnCysGlyLysGlnProAlaLys 520
DB 1834 TTTTCAGATGCCAAGATGTCATCATATATCCCCACACACTGTGGCAAGCAGCCAGCCNA 1893
QY 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
DB 1894 TTTGGGACCATGCTGCTATGTAAGTTGGCGCCAAAGGGTTCAATTTATCTGGAGTCAAAG 1953
QY 541 MetLeuArgCysThrThrSerGlyLysTyrPasnValGlyValGlnAlaValCysLys 560
DB 1954 ATGCTGAGATGTACCACTTCTGGAAAAATGGAATGTGCGAGTTCAGGCAGCTGTGTGTA 2013
QY 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln 580
DB 2014 GAGTGGAGGCTCTCTCAATCACTGTCCTTAAGGACATAGAGGCTAAGGCTCTGGAAACAG 2073
QY 581 GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
DB 2074 CAAGATTCGCCAATGTTACTTGCAGATTCACACAGCTAAAGACAACTCTGGTGAAGAAG 2133
QY 601 ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal 620
DB 2134 GTGTCAGTCCACGTTCACTCAGCTTTCACCCACCTTACCTTTTCCCAATGGAGATGTT 2193
QY 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
DB 2194 GCTATGTAATACGGGCACTGACCTATCCGGCACCGCCAGCTGCATTTTCCATATC 2253
QY 641 LysValIleAspAlaGluProProValIleAspTrpCysArgSerProProValGln 660
DB 2254 AAGTTATTATGTCAGAACCACTCTCATAGACTGGTGCAGATCTCCACCTCCCGTCCAG 2313
QY 661 ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly 680
DB 2314 GTCTCGAGAAAGGTACATGTCGCAAGCTGGGATGAGCTCAGTTCTCAGACAACTCAGGG 2373
QY 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
DB 2374 GCTGAATTTGGTCAATTACCAAGTCATACAGAGGACCTTTTCCCTCAAGGGGAGACT 2433
QY 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
DB 2434 ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATTAACAGGACATGTGATATCCATATT 2493
QY 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
DB 2494 GTCATAAAAGGTTCTCCCTGTGAAATTCATTCACCTGTAAATGGGGATTTTATATGC 2553
QY 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
DB 2554 ACTCCAGATAATAGTGGAGTCAACTGTACATTAATCTTGTGGAGGGCTATGATTTCACA 2613
QY 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTrpLysProThrTyr 780
DB 2614 GAAGGGTCTACTGACAAAGTATTATTGTGCTTATGAAGATGGCGTCTGGAACCAACATAT 2673
QY 781 ThrThrGluTrpProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
```

```
DB 2674 ACCACTGAATGGCCAGACTGTGCCAAAAAACGTTTTTCCAAACCCACGGGTTCCAGTCTCTT 2733
QY 801 GluMetPheTyrLysAlaAlaArgCysAspThrAspLeuMetLysLysPheSerGlu 820
DB 2734 GAGATGTTCTACAAAGCAGCTGCTGTGTATGACACAGATCTGATGAAGAAGTCTTCTGAA 2793
QY 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
DB 2794 GCATTTTGAGAGCGACCCCTGGGAAAAATGGTCCCATCATTTTGTAGTGATGCAGAGGACATT 2853
QY 841 AspCysArgLeuGluGlnAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
DB 2854 GACTGCAGACTGGAGGGAACCTTGACCCAAAAAATTTTGCCTAGAAATATATATGACTAT 2913
QY 861 GluAsnGlyPheAlaIleGly 867
DB 2914 GAAAAATGGCTTTGCAATTGGT 2934
```

Search completed: May 10, 2004, 06:00:17
Job time : 30384 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 10, 2004, 15:48:01 ; Search time 26846.2 Seconds
(without alignments)
3972.182 Million cell updates/sec

Title: US-09-977-053-4
Perfect score: 19973
Sequence: 1 MPRLAFPCNGALVSGWAT.....CHCLSSWYGHNCRKERTGF 3571

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q/cgu2_1/USPTO.spool_p/US09977053/runat 06052004 075943 18200/app query.fasta_1.5710
-DB=EST -OPMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=25 -MODE=LOCAL
-OUTFM=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977053 @CGN 1.1 15412 @runat 06052004 075943 18200 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_man:*
23: em_ges_pro:*
24: em_ges_mus:*
25: em_ges_rtd:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5409.5	27.1	3288	29	AY405591	AY405591 Homo sapi
2	5009.5	25.3	3009	29	AY405592	AY405592 Pan trogl
3	5048	25.3	3127	14	CB548140	CB548140 AWMNNUC:C
4	4803.5	24.0	3288	29	AY405593	AY405593 Mus muscu
5	4781	23.9	3171	11	BC030816	BC030816 Homo sapi
6	4350	21.8	2809	13	BQ832212	BQ832212 C82 AF702
7	3955	19.8	2736	11	AK052699	AK052699 Mus muscu
8	2865.5	14.3	2053	11	AK014693	AK014693 Mus muscu
9	2023.5	10.1	3242	11	AK035333	AK035333 Mus muscu
10	1793	9.0	3307	11	AK052463	AK052463 Mus muscu
11	1690	8.5	906	13	EX390270	EX390270 BX390270
12	1658.5	8.3	1201	13	BX39277	BX39277 BX39277
13	1629.5	8.2	918	14	CD515150	CD515150 AGENCOURT
14	1549	7.8	881	13	BX431419	BX431419 BX431419
15	1470.5	7.4	1201	13	BX356254	BX356254 BX356254
16	1407.5	7.0	1201	13	BX356253	BX356253 BX356253
17	1405.5	7.0	6030	29	AY407073	AY407073 Homo sapi
18	1386.5	6.9	5901	29	AY407075	AY407075 Mus muscu
19	1385	6.9	748	13	AG093982	AG093982 BX093982
20	1382	6.9	769	12	BG534331	BG534331 602553560
21	1362	6.8	940	14	CF407643	CF407643 CH3049 F
22	1359	6.8	738	9	AL706149	AL706149 DKF20686E
23	1319	6.6	839	14	CF409971	CF409971 CH3063 H
24	1304	6.5	801	9	AU137280	AU137280 AU137280
25	1295	6.5	756	12	BG620159	BG620159 602618422
26	1247.5	6.2	3496	11	BC032550	BC032550 Homo sapi
27	1236.5	6.2	792	14	CD109768	CD109768 AGENCOURT
28	1207	6.0	619	13	BU072726	BU072726 in19C07 Y
29	1192	6.0	670	13	CB268566	CB268566 1007472 H
30	1188	5.9	829	13	BG434540	BG434540 602506592
31	1176.5	5.9	701	13	BQ929100	BQ929100 AGENCOURT
32	1145	5.7	719	12	BG436029	BG436029 602508809
33	1128	5.6	626	12	BG535782	BG535782 602563911
34	1125.5	5.6	1887	11	AK045217	AK045217 Mus muscu
35	1124.5	5.6	3348	11	AK003072	AK003072 Mus muscu
36	1113	5.6	900	13	BX454462	BX454462 BX454462
37	1102	5.5	926	13	BQ899350	BQ899350 AGENCOURT
38	1076	5.4	674	29	AG123704	AG123704 Pan trogl
39	1072	5.4	613	10	BE082811	BE082811 RC2-BT064
40	1061	5.3	601	10	BP996491	BP996491 MR2-GN012
41	1057.5	5.3	3809	11	BC032548	BC032548 Homo sapi
42	1050	5.3	2408	11	AK087058	AK087058 Mus muscu
43	1046	5.2	9330	11	AK090118	AK090118 Mus muscu
44	1039.5	5.2	2019	11	BC047729	BC047729 Homo sapi
45	1030	5.2	653	9	AV610969	AV610969 AV610969
46	1027.5	5.1	1035	12	BG621817	BG621817 602619678
47	1027	5.1	564	14	CP903215	CP903215 A0402E03-
48	1023	5.1	702	14	CB521378	CB521378 UI-M-GH0-
49	1022	5.1	642	10	BB500305	BB500305 BB500305
50	1016	5.1	969	28	CC266833	CC266833 CH261-124
51	1013	5.1	608	10	BG010666	BG010666 PM3-GN018
52	1012	5.1	651	14	CB434567	CB434567 611194 MA
53	1011	5.1	548	14	CB269729	CB269729 1008636 H
54	998	5.0	664	14	CD352234	CD352234 UI-M-GH0-
55	996	5.0	701	12	BG966366	BG966366 602832856
56	990	5.0	673	10	BB542219	BB542219 BB542219
57	987	4.9	657	10	BB626617	BB626617 BB626617
58	987	4.9	789	13	BU624529	BU624529 UI-H-FG1-
59	982.5	4.9	998	13	BQ717243	BQ717243 AGENCOURT
60	982	4.9	763	13	BQ392790	BQ392790 603805162
61	982	4.9	964	12	BG243790	BG243790 602357810
62	969.5	4.9	656	13	BY714190	BY714190 BY714190

63	967	4.8	553	14	N77999	N77999 yv70g02.x1
64	958	4.8	652	13	BQ367139	BQ367139 MR3-GN018
65	956	4.8	543	14	CD693738	CD693738 EST10261
66	955	4.8	526	13	BQ558339	BQ558339 H4051F02-
67	948	4.7	567	10	BG012967	BG012967 IL5-GN023
68	943	4.7	496	9	AA419490	AA419490 zv03a02.r
69	937	4.7	512	14	CF903931	CF903931 A0415G09-
70	929	4.7	563	14	R71349	R71349 Yf54e12.x1
71	928	4.6	526	10	BG012979	BG012979 IL5-GN023
72	926	4.6	599	9	AV594654	AV594654 AV594654
73	922	4.6	503	10	BF367968	BF367968 RC3-GN004
74	922	4.6	907	12	BG435117	BG435117 602507128
75	921	4.6	5824	29	AY407074	AY407074 Pan trogl
76	916	4.6	709	29	AG124019	AG124019 Pan trogl
77	915	4.6	508	14	CA540087	CA540087 C0286G08-
78	908	4.5	606	14	CD766382	CD766382 AGENCOURT
79	898	4.5	482	10	BF998780	BF998780 QV0-GN023
80	892	4.5	946	14	CP407642	CP407642 CH3#049.F
81	890	4.5	474	12	EM313719	EM313719 ih05e04.x
82	888	4.4	476	12	BM313989	BM313989 ih05e04.y
83	884	4.4	487	14	N91089	N91089 za18c09.r1
84	884	4.4	569	10	BG012978	BG012978 IL5-GN023
85	883	4.4	551	12	BF535506	BF535506 399051.MA
86	878	4.4	574	10	BE810281	BE810281 RC3-HN000
87	876	4.4	543	12	BI340183	BI340183 365377.MA
88	864	4.3	489	10	BG012946	BG012946 IL5-GN023
89	857	4.3	576	10	AW983663	AW983663 RC3-HN000
90	856	4.3	480	10	BG003364	BG003364 RC3-GN004
91	853	4.3	705	12	BM972572	BM972572 UI-CF-EC1
92	852	4.3	696	12	BI453395	BI453395 603174360
93	851	4.3	449	9	AA424657	AA424657 zv91e07.r
94	850	4.3	445	10	BF108713	BF108713 7145a01.x
95	846	4.2	461	10	BF986530	BF986530 CM3-GN005
96	843	4.2	579	10	BF367305	BF367305 MR0-GN002
97	839	4.2	905	12	BG174788	BG174788 602336052
98	837	4.2	530	12	BM705310	BM705310 UI-E-C11-
99	831	4.2	461	12	EM721564	EM721564 UI-E-E01-
100	830	4.2	619	9	AA402528	AA402528 zu48c06.r

RESULT 1	AY405591	3288 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	Homo sapiens HCM2275 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
DEFINITION	AY405591				
VERSION	AY405591.1	GI:39761565			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 3288)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,P., Murphy,B., Ferrieca,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 3288)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,P., Murphy,B., Ferrieca,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				

FEATURES	Location/Qualifiers
source	1..3288
gene	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	<1..>3288
ORIGIN	/locus_tag="HCM2275"
Alignment Scores:	
Pred. No.:	0 Length: 3288
Score:	5409.50 Matches: 1000
Percent Similarity:	90.42% Conservatives: 1
Best Local Similarity:	90.33% Mismatches: 87
Query Match:	27.08% Indels: 19
DB:	29 Gaps: 1
US-09-977-053-4 (1-3571) x AY405591 (1-3288)	
Qy	1 MetTTPProArgLeuAlaPheCysTyrTrpGlyLeuAlaLeuValSerGlyTTPAlaThr 20
Db	1 ATGTGGCCTCGCCTGGCCTTTTGTTCCTGGGGTCTGGCGCTCGTTTCGGGCTGGGGACC 60
Qy	21 PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
Db	61 TTTCAGCAGATGTCCCGTGGCGCAATTTTCAGCTTCCCGCTNNNNNNNNNNNNNNNN 120
Qy	41 GlyAlaProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
Db	121 NNN 180
Qy	61 ValGluArgLeuGlyClnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
Db	181 NNN 240
Qy	81 ArgLeuGluLeuValPheLeuValAspSerSerValGlyGluValAsnPheArg 100
Db	241 CCCTGGAGCTTGTCTTCTTGTGGTGAATTCGTCCAGCGTGGCGGCAAGTCAACTCCGC 300
Qy	101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
Db	301 AGCAGACTCATGTCCTCGCAAGCTGTCTCCGACTTCCCGTGGTGGCCACGGCCACG 360
Qy	121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
Db	361 CGCGTGGCCATCGTGACCTTCTCGTCCAGAACTACGTGGTGGCGCGTGCATTCATC 420
Qy	141 SerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluLeuProAlaIle 160
Db	421 TCCACCCCGCGCGCGCCAGCAGTCCGCGTCTCTCCAGAGATCCCTGCGCATC 480
Qy	161 SerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPheGlnGlnAlaAlaGlnIleLeu 180
Db	481 TCTTACCGAGTGGCGGCACCTACCAAGGGGGCTTCCAGCAAGCCGCAAAATCTT 540
Qy	181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerArg 200
Db	541 CTTTCATGCTAGAGAAAACCTCAACAAAAGTTGATTTCTCATCTCATGATGATATTCAT 600
Qy	201 GlyGlyAspProArgProIleAlaIleSerLeuArgAspSerGlyValGluIlePheThr 220
Db	601 GGGGAGAGCCCTAGACCATTGCGGTGCTGCGGAGATTCAGGATTCAGGATTCCTCACT 660
Qy	221 PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu 240
Db	661 TTTCGCATATGCAAGGAAACATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGAG 720
Qy	241 GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLeu 260
Db	721 GAGCACTGTGTACCTGCTACACAGTTTGAAGAAATTTAGGGCTTTAGCTCGCGGGCATG 780
Qy	261 HisGluAspLeuProSerGlySerPheIleGlnAspMetValHisCysSerTyrLeu 280

Db 781 CATGAAGATCTACCTCTCGGAGTGTATTAACAGATGATATGGTCCACTCATATCTT 840
Qy 281 CysAspGluGlyAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr 300
Db 841 TGTGATGAAGGCAAGGACTGCTGTGACCGAATGGGAAGCTGCAAAATGTGGGACACACA 900
Qy 301 GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys 320
Db 901 GGCCATTTGAGTGCATCTGTGAAGAAGGGTATTACGGGAAGGCTGTGCAATGATGATGC 960
Qy 321 ThrAlaCysProSerGlyThrTyrLysProGluGlySerProGlyGlyIleSerSerCys 340
Db 961 ACAGCTTGCCTCCATCGGGGACATCAAAACCTGGAAGGCTCACAGAGGAATCAGCAGTTGC 1020
Qy 341 IleProCysProAspGluHisThrSerProGlySerThrSerProGluAspCys 360
Db 1021 ATTCCATGTCCTGATGAATAATCACTCTCCACCTTGGGAAGCAGCATCCCTGGAAGACTGT 1080
Qy 361 ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla 380
Db 1081 GTCTGCAGAGGGAATACAGGGCATCTGGCCAGACTGTGAACTTGTCCACTGCCCTGCC 1140
Qy 381 LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla 400
Db 1141 CTGAAGCCTCCCGAAATGGTTACTTTATCCAAAACACTTGCAACCAACCACTTCAATGCA 1200
Qy 401 AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleIleLeuCys 420
Db 1201 GCGTGTGGGGTCCGATGTCACTCCGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT 1260
Qy 421 LeuProAsnGlyLeuTyrSerGlySerGluSerTyrCysArgValArgThrCysProHis 440
Db 1261 CTACCCAAATGGTTGTGGTTCGGTTCAGAGACTACTGCAAGATAGACATGTCTCAT 1320
Qy 441 LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr 460
Db 1321 CTCGGCAGCGAAACATGGCCACATCAGCTGTCTACAGGGAATGTTATATAAGACA 1380
Qy 461 ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys 480
Db 1381 ACATGTTGTGGTTCCTGTGATGAAGGGTACAGACTAGAAGGCGCATGATAGCTTACTTGT 1440
Qy 481 GlnGlyAsnSerGlnTyrAspGlyProGluProArgCysValGluArgHisCysSerThr 500
Db 1441 CAAGGAAACAGCCAGTGGGAATGGCCCAAGAACCCCGGTGTGTGGAGGCCACTGTTCCACC 1500
Qy 501 PheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAlaLys 520
Db 1501 TTTTCAGATGCCCAAGATGTCATCATATATCCCCCAACACTGTGCAAGCAGCCAGCCAAA 1560
Qy 521 PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu 540
Db 1561 TTTGGACCATCTGCTATGTAAGTTCGCCCAAGGGTTCATTTTATCTGGAGTCAAGAA 1620
Qy 541 MetLeuArgCysThrThrSerGlyLysTyrAsnValGlyValGlnAlaValCysLys 560
Db 1621 ATGCTGAGATGTACCACTCTCGAANAATGGAAATGTGAGTTCAGGAGCTGTGTGTAAA 1680
Qy 561 AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGln 580
Db 1681 GACGTGGAGGCTCTCAAAATCACTGCTCTTAAGGACATAGAGGCTAAGACTCTGGAACAG 1740
Qy 581 GlnAspSerAlaAsnValThrTyrGlnIleProThrAlaLysAspAsnSerGlyGluLys 600
Db 1741 CAAGATTTCTGCCAATGTATCTGGCGAGATTCCCAACAGCTTAAAGACAACTCTGGTGAAG 1800
Qy 601 ValSerValHisValHisProAlaPheThrProTyrIleuPheProIleGlyAspVal 620
Db 1801 GTGTCACTCACTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGTT 1860
Qy 621 AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle 640
Db 1861 GCTATCGTATACAGGCAACTGACCTATCCGGCAACCAAGGCCAGCTGCGTTTTCATATC 1920

Qy 641 LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln 660
Db 1921 AAGTATTATGATGCAAGAACCACTGTCTATAGACTGGTGCAATCTCCACCTCCCGTCCAG 1980
Qy 661 ValSerGluLysValHisAlaAlaSerTyrAspGluProGlnPheSerAspAsnSerGly 680
Db 1981 GTCTCGGAGAAGGTACATGCGCAAGCTGGGATGAGCTCAGTTCTTCAGAACATCAGGG 2040
Qy 681 AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr 700
Db 2041 GCTCAATTTGCTCATTTACAGAGTCAATACACAGGAGACCTTTTCTCCCTCAAGGGGAG 2100
Qy 701 IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle 720
Db 2101 ATAGTACGTATATACGCACTGACCCCTCAGGCAATACAGGACATGTATATCATATT 2160
Qy 721 ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys 740
Db 2161 GTCATAAA----- 2169
Qy 741 ThrProAspAsnThrGlyValAsnCysThrLeuThrCysLeuGluGlyTyrAspPheThr 760
Db 2170 -----GATAATATCTGGAGTCAACTGTACATTAATCTGCTTGGAGGCTATGATTTCAC 2223
Qy 761 GluGlySerThrAspLysTyrTyrCysAlaTyrGluAspGlyValTyrLysProThrTyr 780
Db 2224 GAAGGCTTACTGACAGATTAATTGTGCTTATGAGATGGCGCTCTGGAAACCAACATAT 2283
Qy 781 ThrThrGluTyrProAspCysAlaLysLysArgPheAlaAsnHisGlyPheLysSerPhe 800
Db 2284 ACCACTGAATGGCCAGACTGTGCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2343
Qy 801 GluMetPheTyrLysAlaAlaArgCysAspAspThrAspLeuMetLysLysPheSerGlu 820
Db 2344 NNN 2403
Qy 821 AlaPheGluThrThrLeuGlyLysMetValProSerPheCysSerAspAlaGluAspIle 840
Db 2404 NNN 2463
Qy 841 AspCysArgLeuGluAsnLeuThrLysLysTyrCysLeuGluTyrAsnTyrAspTyr 860
Db 2464 GACTGCAGACTGCGAGGAGAACCTTGACCAAAAAATATTGCTTAGAATATAATTATGACTAT 2523
Qy 861 GluAsnGlyPheAlaIleGlyProGlyTyrGlyAlaAlaAsnArgLeuAspTyrSer 880
Db 2524 GAAAAATGGCTTTTGCAATTGGACAGGTGGCTGGGGTGCAGCTAATAGGCTGGATTACTCT 2583
Qy 881 TyrAspAspPheLeuAspThrValGlnGluThrAlaThrSerIleGlyAsnAlaLysSer 900
Db 2584 TAGATGACTTCTTGGACACTGTGCAGAAACAGCCACAGCATCGGCNAATGCCAAGTCC 2643
Qy 901 SerArgIleLysArgSerAlaProLeuSerAspTyrLysIleLysLeuIlePheAsnIle 920
Db 2644 TCACGGAATTAAGAGAGTGGCCCAATATCTGACTATATAAATTAAGTTAAATTTTAAACATC 2703
Qy 921 ThrAlaSerValProLeuProAspGluArgAsnAspThrLeuGluTyrGluAsnGln 940
Db 2704 ACAGCTAGTGTGGCATTTACCCGATGAAGAAATGATACCTTGAATGGGAATTCAGCAA 2763
Qy 941 ArgLeuLeuGlnThrLeuGluThrIleThrAsnLysLeuLysArgThrLeuAsnLysAsp 960
Db 2764 CGACTCTTCAGACATTTGGAAACTATACAAATAAATGAAGAGACTCTCACAAAGAC 2823
Qy 961 ProMetTyrSerPheGlnLeuAlaSerGluIleLeuIleAlaAspSerAsnSerLeuGlu 980
Db 2824 CCCATGTATTCTTTCAGCTTGCATCAGAAATATCTATAGCCGACAGCAATTCATTAGAA 2883
Qy 981 ThrLysLysAlaSerProPheCysArgProGlySerValLeuArgGlyArgMetCysVal 1000
Db 2884 AAAAAAAGGCTTCCCCCTTCTCGCAGACCGAGCTCAGTGTGTGAGAGGGCGTATGTGTGTC 2943

```

QY 1001 AsnCysProLeuGlyThrTyrTyrAsnLeuGluHisPheThrCysGluSerCysArgIle 1020
DB 2944 AATTCGCCCTTTGGGAACCTATTATAATCTGGAAACATTTTCACCTGTGAAAGCTGCCGGATC 3003
QY 1021 GlySerTyrGlnAspGluGlyGlnLeuGluCysLysLeuCysProSerGlyMetTyr 1040
DB 3004 GGATCTTATCAGATGAGAGAGGACACTTGAGTGCAGCTTTGGCCCTCTGGGATGAC 3063
QY 1041 ThrGluTyrIleHisSerArgAsnIleSerAspCysLysAlaGlnCysLysGlnGlyThr 1060
DB 3064 ACGGAATATATCCATTCAGAAACATCTCTGATTTGTAAGCTCAGTGTAAACAAGGCACC 3123
QY 1061 TyrSerTyrSerGlyLeuGluThrCysGluSerCysProLeuGlyThrTyrGlnProLys 1080
DB 3124 TACTCATACAGTGGACTTGAGACTTGTGAATCGTGTCCACTGGGCACCTTATCAGCCAAA 3183
QY 1081 PheGlySerArgSerCysLeuSerCysProGluAsnThrSerThrValLysArgGlyAla 1100
DB 3184 TTTGGTTCCGGAGTGCTCTCTGTGTCCAGAAACACCTCAACTGTGAAAGAGAGGCC 3243
QY 1101 ValAsnIleSerAlaCysGly 1107
DB 3244 GTGAACATTTCTGCATGTGA 3264

RESULT 2
AY405592 3009 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM2275 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405592
VERSION AY405592.1 GI:39761566
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 3009)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3009)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source
        1..3009
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
    gene
        <1..>3009
            /locus_tag="HCM2275"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3009
Score: 5060.50 Matches: 928
Percent Similarity: 91.81% Conservative: 3
Best Local Similarity: 91.52% Mismatches: 64
Query Match: 25.34% Indels: 19
DB: 29 Gaps: 1

US-09-977-053-4 (1-3571) x AY405592 (1-3009)

```

```

QY 94 ValGlyGluValAsnPheArgSerGluLeuMetPheValArgLysLeuLeuSerAspPhe 113
DB 1 GTGGGCGAAGTCAACTTCGCGAGCGAGCTCATGTTCGTCGCGAAGCTGCTGCCGACTTC 60
QY 114 ProValValProThrAlaThrArgValAlaIleValThrPheSerSerLysAsnTyrVal 133
DB 61 CCGGTGTGCCACGCGCCACGCGGTGGCCATCTGTGACCTTCCTGTCGAGAGACTACGTG 120
QY 134 ValProArgValAspTyrIleSerThrArgArgAlaArgGlnHisLysCysAlaLeuLeu 153
DB 121 GTGGCGGCGGTGCATTTACATCTCCACCGCGCGCGCGCAGCACAAAGTCGCGCTGCTC 180
QY 154 LeuGlnGluIleProAlaIleSerTyrArgGlyGlyGlyThrTyrThrLysGlyAlaPhe 173
DB 181 CTCGAAGAGATCCCTGCCATCTCTACCGAGGTGGCGCACCTACACCAAGGCGCGCTTC 240
QY 174 GlnGlnAlaAlaGlnIleLeuLeuHisAlaArgGluAsnSerThrLysValValPheLeu 193
DB 241 CAGCAAGCGCGCAAAATTTCTTCTCATGNNANNGAANAACCTCAACAAGATTATATTCTC 300
QY 194 IleThrAspGlyTyrSerAsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAsp 213
DB 301 ATCACTGATGGATATTCCAAATGGCGGNNACCTTAGACCAATTCGCGCTCACTGNNAGAT 360
QY 214 SerGlyValGluIlePheThrPheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAsp 233
DB 361 TCAGNNNTGGAGATCTTCACTTTTGGCATATGGCAAGGGAACATTGAGAGCTGGAATGAC 420
QY 234 MetAlaSerThrProLysGluGluHisCysTyrLeuLeuHisSerPheGluGluPheGlu 253
DB 421 ATGGCTTCCACCCCAAGNNNNNNCTGTGTACTCTGTACACAGTTTGTGANGAATTGTAN 480
QY 254 AlaLeuAlaArgArgAlaLeuHisGluAspLeuProSerGlySerPheIleGlnAspAsp 273
DB 481 NCTTTAGCTCGCGCGGCATTGCGATGANNATCTACCTTCTGGGAGTTTATTTCAGATGAT 540
QY 274 MetValHisCysSerTyrLeuCysAspGluGlyLysAspCysCysAspArgMetGlySer 293
DB 541 ATGGTCCACTGCTCATATCTTTGTGATGAAGGACAGACTGCTGTGACCGAATGGCAGC 600
QY 294 CysLysCysGlyThrHisThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGly 313
DB 601 TGCAAATGTGGACACACACAGCCCATTTTGATGTGATCTGTGAAGAAGGGTATTACGGG 660
QY 314 LysGlyLeuGlnTyrGluCysThrAlaCysProSerGlyThrTyrLysProGluGlySer 333
DB 661 AAAGGTCTGCAGTATGATGACAGCTTGCCCATCGGGGACATACAAACCTGGAAGGCTCA 720
QY 334 ProGlyGlyIleSerSerCysIleProCysProAspGluAsnHisThrSerProGly 353
DB 721 CCAGGAGGAATCAGCAGTTGCATTCATGCTCTGATGAANAATCACACCTCTCCACCTGGA 780
QY 354 SerThrSerProGluAspCysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCys 373
DB 781 ACCACATCCCTGAGAGACTGTGTCTGCAGAGAGGATACAGGCACTCTGGCCAGACCTGT 840
QY 374 GluLeuValHisCysProAlaLeuLysProProGluAsnGlyTyrPheIleGlnAsnThr 393
DB 841 GAACCTGTCCACTGCGCTGCGCTGGAAGCTCCGAAAAATGGTTACTTTATATCCAAAACT 900
QY 394 CysAsnAsnHisPheAsnAlaAlaCysGlyValArgCysHisProGlyPheAspLeuVal 413
DB 901 TGCACCAACCACTTCAATGACGCTGTGGGGTCCGATGTACCCTGGATTGTATCTGTG 960
QY 414 GlySerSerIleIleLeuCysLeuProAsnGlyLeuTrpSerGlySerGluSerTyrCys 433
DB 961 GGAAGCAGCACTATCTATGTCTACCCATGTTGTGTGTGTGCTCCGTTCCAGAGACTACTGC 1020
QY 434 ArgValArgThrCysProHisLeuArgGlnProLysHisGlyHisIleSerCysSerThr 453
DB 1021 AGAGTAAGAACAATGTCTCATCTCCGCGCAGCCGAAACATGGCCACATCAGCTGTTCTACA 1080

```


454 ArgGluMetLeuTyrLysThrThrCysLeuValAlaCysAspGluGlyTyrArgLeuGlu 473
Db AGGAAATGTTATATAAGACACATCTTTGGTGGCTGTGATGAGGGTACAGCTAGAA 1140
474 GlySerAspLysLeuThrCysGlnGlyAasnSerGlnTrpAspGlyProGluProArgCys 493
Db GGCAGTGATAAGCTTACTTGTCAAGGAAACAGCCAGTGGGATGGGCGCAGAACCCCGTGT 1200
494 ValGluArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisAasn 513
Db GTGGAGCGCCAGTGTTCACCTTTTCAGACCCCAAGAGTGCATCATATCCCCCCCAAC 1260
514 CysGlyLysGlnProAlaLysPheGlyThrIleCysGlyTyrValSerCysArgGlnGlyPhe 533
Db TGTGGCAGCAGCCAGCCAAATTTGGACGATCTGCTATGTAAGTTGCCCGCAGGGTTC 1320
534 IleLeuSerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTrpAasnValGly 553
Db ATTTTATCTGGAGTCAAGAAATGCTCAGATGTACCACCTTCGGAAATGGAAATGCGGA 1380
554 ValGlnAlaAlaValCysLysAspValGluAProGlnIleAasnCysProLysAspIle 573
Db GTTCAGGCGAGCTGTGTGTAAGNNNNNNNNNNCTCTCAAAATCAACTGCTCTAAGGACATA 1440
574 GluAlaLysThrLeuGluGlnAaspSerAlaAasnValThrTrpGlnIleProThrAla 593
Db GNGNCTAAGACTCTGGACAGCAAGATTTGCGCAATGTTACCTGGCAGATTCACACGCT 1500
594 LysAaspAasnSerGlyGluLysValSerValHisValHisProAlaPheThrProProTyr 613
Db AAAGACAACTCTGGTGAAGGTTGTCAGTCCAGCTTCATCCAGCTTTCACCCCACTTAC 1560
614 LeuPheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeuSerGlyAasnGln 633
Db CTTTTCCTCCCAATTTGGAGATGTGTATCTGATATACACGGCAACTCACTATCCGCAACCCAG 1620
634 AlaSerCysIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCys 653
Db GCGAGCTGGGTTTCCATATCAAGGTATTGATGCGAGAACCACTGTCTAGACTGGTGC 1680
654 ArgSerProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluPro 673
Db AGATCTCCACCTCCCGTCCAGGCTCTCAGAGAGGTATCATGCTGCAAGCTGGGATGAGCCT 1740
674 GlnPheSerAspAasnSerGlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAsp 693
Db CAGTTCTCAGACAACCTCAGCGGGCTGAATTTGGTCAITACCAAGAGTATATACCAAGGAGAC 1800
694 LeuPheProGlnGlyGluThrIleValGlnTyrThrAlaThrAspProSerGlyAasnAasn 713
Db CTTTTCCTCAAGGGAGACTATAGTACAGTATACAGCCACTGACCCCTCAGGCATATAC 1860
714 ArgThrCysAspIleHisIleValIleLysGlySerProCysGluIleProPheThrPro 733
Db AGGACATGTGATATCATATTGTCTATAAAA----- 1890
734 ValAasnGlyAspPheIleCysThrProAaspAasnThrGlyValAasnCysThrLeuThrCys 753
Db -----GATAATACTGGAGTCAACTGTATCAATTAACTTGC 1923
754 LeuGluGlyTyrAspPheThrGluGlySerThrAspLysTyrTyrCysAlaTyrGluAasp 773
Db TTGGAGGGCTATGATTTTCAGAGAGGGTCTACTGACAGATTAATTATTTGCTTATGAGAT 1983
774 GlyValTrpLysProThrTyrThrThrGluTrpProAspCysAlaLysLysArgPheAla 793
Db GCGCTCTGGAACCAACATATACCTGTAATGCCAGACTGTGCCANNNNNNNNNNNN 2043
794 AasnHisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCysAspAspThrAsp 813
Db NNN 2103
814 LeuMetLysPheSerGluAlaPheGluThrLeuGlyLysMetValProSerPhe 833

2104 NNN 2163
834 CysSerAspAlaGluAaspIleAaspCysArgLeuGluGluAasnLeuThrLysLysTyrCys 853
Db TGTAGTATGATGACAGGACATTTGACTGCAGACTGGAGGAGAACCTGCACCAAAAATATTTC 2223
854 LeuGluTyrAasnTyrAspTyrGluAasnGlyPheAlaIleGlyProGlyGlyTrpGlyAla 873
Db CTAGATATATATATGACTATGAAATGGCTTTGCAATTGGACAGGTGGCTGGGGTGA 2283
874 AlaAasnArgLeuAaspTyrSerTyrAspAaspPheLeuAaspThrValGlnGluThrAlaThr 893
Db GCTAATAGGCTGGATTAATCTTATGATGACTTCTCTGGACACTGTGCAAGAAACACCCACA 2343
894 SerIleGlyAasnAlaLysSerSerArgIleLysArgSerAlaProLeuSerAspTyrLys 913
Db AGCATCGCAATGCGAGTCAATCGGATTAAGAAAGTGGCCCATTAATCTGACTATATA 2403
914 IleLysLeuIlePheAasnIleThrAlaSerValProLeuProAaspGluArgAasnAspThr 933
Db ATTAAGTTAATTTTAAATCAATCAGCTAGTGTGCCATTAACCGATGAAAGAAATGATACC 2463
934 LeuGluTrpGluAasnGlnArgLeuLeuGlnThrLeuGluThrIleThrAasnLysLeu 953
Db CTTGAATGGGAAATCAGCAACGACTCTCTCAGACATTTGGAACCTATCACAATAAATCTG 2523
954 LysArgThrLeuAasnLysAaspProMetTyrSerPheGlnLeuAlaSerGluIleLeuIle 973
Db ANAAGGACTCTCAACAAAGACCCCAAGTATTTCTTTTTCAGCTTGCATCAGAAATCTTATA 2583
974 AlaAaspSerAasnSerLeuGluThrLysLysAlaSerProPheCysArgProGlySerVal 993
Db GCGACAGCAATTCATTTAGAAACAAAGAAAGGTTCCCTCTTCGACAGACAGGCTCAGTG 2643
994 LeuArgGlyArgMetCysValAasnCysProLeuGlyThrTyrTyrAasnLeuGluHisPhe 1013
Db CTGAGAGGGGCTATGTGTCAATTTGGTCCCTTTGGGAACCTATTATTAATCTGGAACATTTTC 2703
1014 ThrCysGluSerCysArgIleGlySerTyrGlnAaspGluGlyGlnLeuGluCysLys 1033
Db ACCTGTGAAGCTGCCCGATCGGATCTCTCAAGATGAAGAGGCAACTTTGAGTGCAG 2763
1034 LeuCysProSerGlyMetTyrThrGluTyrIleHisSerArgAasnIleSerAspCysLys 1053
Db CTTTGGCCCTCTGGATGTACCGAATATATCCATTCAGAAACATCTCTGATTGTATA 2823
1054 AlaGlnCysLysGlnGlyThrTyrSerTyrSerGlyLeuGluThrCysGluSerCysPro 1073
Db GCTCAGTGTAAACCAAGCCACTACTCATACAGTGGACTTGAGACTTGTGAATCGTGTCCA 2883
1074 LeuGlyThrTyrGlnProLysPheGlySerArgSerCysLeuSerCysProGluAasnThr 1093
Db CTGGGCACTTATCAGCCAAAATTTGGTTCCGGAGCTGCCCTCTCGTGTCCAGAAACACC 2943
1094 SerThrValLysArgGlyAlaValAasnIleSerAlaCysGly 1107
Db TCACCTGTGAAGAGAGCGCGTGAACATTTCCNNATGTGGA 2985
2944

RESULT 3
CB548140
LOCUS
DEFINITION
AMGNNUC.CRI-00144-A5-Y Colon Rat 1 (10390) Rattus norvegicus cdna
clone cri-00144-a5, mRNA sequence.
ACCESSION
CB548140
VERSION
CB548140.1
KEYWORDS
GI:29432081
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 3127)

1808 GlyAlaValThrPheSerCysGlnGluGlyTyrGlnLeuMetGlyValThrLysile 1827
 1742 GGTGCTGAAGTACATTTCTCTGTGAAGAGGACCAGCTGGTGGAGTGAAGAAATC 1801
 1828 ThrCysLeuGlnSerGlyGluThrPheHisLeuLeuProTyrCysLysAlaValSerCys 1847
 1802 ACCTGTTTGGAGTCTGGCGAGTGGATACCTCAGACCATCTCTGTGAAGCCATTTCTCTGT 1861
 1848 GlyLysProAlaIleProGluLysAsnGlyCysIleGluGluLeuAlaPheThrPheGlySer 1867
 1862 GGTGCCCCACCTGTTCTCTGAAATGGCGGTGTGACCGGTGAGCTTTCATCATATGGCAGT 1921
 1868 LysValThrTyrArgCysAsnLysGlyTyrThrLeuAlaGlyAspLysGluSerCys 1887
 1922 AAGGTGAGGTACAGGTGTGATAAAGCTATACCTTGGCCGGGATGAAGAGTCAGCATGC 1981
 1888 LeuAlaAsnSerSerTrpSerHisSerProProValCysGluProValLysCysSerSer 1907
 1982 CTTGCTAGTGGTTCGTGGAGTCACTCTCTCTCTGTGTGTGAGCTGGTGAAGTGTTCCTCAG 2041
 1908 ProGluAsnIleAsnGlnGlyLysTyrIleLeuSerGlyLysLeuThrTyrLeuSerThrAla 1927
 2042 CTTGAGAACATAATACGGGAATACATCTTAAGCGGGTTCACCTACCTTTTCATGTGCA 2101
 1928 SerTyrSerCysAspThrGlyTyrSerLeuGlnGlyProSerIleIleGluCysThrAla 1947
 2102 TCATATTTCTGTGAGGAGGATACAGCTTACAGGCGCGTGGTCTCATTTGAATGCACAGCT 2161
 1948 SerGlyIleTrpAspArgAlaProProAlaCysHisLeuValPheCysGlyGluProPro 1967
 2162 TCCGGCAGCTGGGACAGAGCCACCTAGTTGTCACTTGTGTCTCTCGGAGAGCCTCCA 2221
 1968 AlaIleLysAspAlaValIleThrGlyAsnAsnPheThrPheArgAsnThrValThrTyr 1987
 2222 ATGGTCAAGATGCTCCACACAGGAGGCACTTCACTTTTGGAAACATGCTCACTTAC 2281
 1988 ThrCysLysGluGlyTyrThrLeuAlaGlyLeuAspThrIleGluCysLeuAlaAspGly 2007
 2282 ACATGACAGGAGGCTACACCTTGGCGGCTGACACCATCATATGCCAGGCCAATGGC 2341
 2008 LysTrpSerArgSerAspGlnGlnCysLeuAlaValSerCysAspGluProProIleVal 2027
 2342 AAGTGGAAATCAAGTAAACACAGTGTCTGGCGGTCTCTCGGACGAGCCGCCCAATGTG 2401
 2028 AspHisAlaSerProGluThrAlaHisArgLeuPheGlyValAspIleAlaPheTyrTyrCys 2047
 2402 GACACGCTCTCCAGAGATGGCGACAGAGCTCTTGGGGACACCGGTTCTACTACTGT 2461
 2048 SerAspGlyTyrSerLeuAlaAspAsnSerGlnLeuLeuCysAsnAlaGlnGlyLysTrp 2067
 2462 GCAGATGGTTACAGCTGGCTGGACAAATTCACAGTCACTGCAATGCCAGGGGAACCTGG 2521
 2068 ValProProGluGlnAspMetProArgCysIleAlaHisPheCysGluLysProPro 2087
 2522 GTTCCCGCCGAGGCGCAGGCTGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2581
 2088 SerValSerTyrSerIleLeuGluSerValSerLysAlaLysPheAlaGlySerVal 2107
 2582 TCTGTTTCTTACAGCATCTTGGAGTCTGTGAGCAAGCAAAATTTGAGCTGGCTCGGTA 2641
 2108 ValSerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMet 2127
 2642 GTAAAGTTCAAGTGCATGGAGGTTTCGTACTGAACACCTCAGCGAAGATTGAATGGCTG 2701
 2128 ArgGlyGlyGlnTrpAsnProSerProMetSerIleGlnCysIleProValArgCysGly 2147
 2702 AGAGGCGGACAGTGGAGCCCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2761
 2148 GluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyrSerPheGlyAlaMet 2167
 2762 GAGCCTTCAAGCATCAAAATGGCTACCCGAGTGGAAACGAATACAGTTTTCGGGGCGGTG 2821
 2168 ValAlaTyrSerCysAsnLysGlyPheTyrIleLysGlyGluLysLysSerThrCysGlu 2187

2822 GTGCGTACAGCTGCCACAAGGGATTCTATATCAAAAGGGAGAAAGACGACATGTGAG 2881
 2188 AlaThrGlyGlnTrpSerSerProIleProThrCysHisProValSerCysGlyGluPro 2207
 2882 GCCACGGGACAGTGGAGTAGACCTCTGCCACCTGCGCACCGGTATCTCTGAACGACCA 2941
 2208 ProLysValGluAsnGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluVal 2227
 2942 CCTAAGGTTGAGAACGGCTTCTCTGGAGCACACCACTGGCAGGACCTTTGAGAGCGAAGCA 3001
 2228 ArgTyrGlnCysAsnProGlyTyrLysSerValGlySerProValPheValCysGlnAla 2247
 3002 AGGTTCCAGTGCACCCAGGTTACAGGACAGTTCGAGAGCCCTGTGTTTGTCTGCGCAAGCC 3061
 2248 AsnArgHisThrHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysPro 2267
 3062 AACGACACTGGACAGTGAAGCCCTCTGTCTGACGCTCTCACTGTGGGAAACCC 3121
 2268 ProPro 2269
 3122 CCCCC 3127
 RESULT 4
 AY405593
 LOCUS Mus musculus HCM2275 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 ACCESSION AY405593
 VERSION AY405593.1 GI:39761567
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 3288)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3288)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source 1..3288
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..3288
 /locus_tag="HCM2275"
 gene
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3288
 Score: 4803.50 Matches: 879
 Percent Similarity: 84.74% Conservative: 76
 Best Local Similarity: 77.99% Mismatches: 139
 Query Match: 24.05% Indels: 33
 DB: 29 Gaps: 5
 US-09-977-053-4 (1-3571) x AY405593 (1-3288)

QY 1 MetTrpProArgLeuAlaPheCysCysTrpGlyLeuAlaLeuValSerGlyTrpAlaThr 20
DB 1 ATGTGTCGCGCTGGCCCTTTGTTGCTGGGCTCTGGCACTGGTGTCAAGCTGGACCAAC 60
QY 21 PheGlnGlnMetSerProSerArgAnPheSerPheArgLeuPheProGlnThrAlaPro 40
DB 21 PheGlnGlnMetSerProSerArgAnPheSerPheArgLeuPheProGlnThrAlaPro 40
DB 61 TTCAGCGCTGGCCCTTGGCTCAACTTCAGCTTCGGCTGTTCGCCGAGGCTCTCCG 120
QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
DB 121 GGGCTCTGGCAGACTGGCGGTACCTCCGCGTCCAGTGAGGAGGAGCAGCAGGAGC 180
QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgValArgValArgLeuLeuArgGluLeuSer 79
DB 181 AAAGTGGAGCGCTGGCGCGCGCTTCCGAGCGCGTGGCGGAGCTGGCGGAGCTCAGC 240
QY 80 GluArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAlaAnPhe 99
DB 241 GGCAGCTGGAGCTCTCTCTGGTGGAGAGTCTGTCAGCGTGGCGGCAACCAACTTC 300
QY 100 ArgSerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAla 119
DB 301 CTCACAGAGCTCAAGTCTGGTGGCAAGCTGCTGTCGACTTCCCGTGGTGTCCACGGCC 360
QY 120 ThrArgValAlaIleValThrPheSerSerLysAsnTrpValValProArgValAspTyr 139
DB 361 ACCGTGTGGCCATCTGCTCACTTCTCATCCAGAACACAGTGTGGTGGCGGCTGAGTAC 420
QY 140 IleSerThrArgArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAla 159
DB 421 ATCTCACAGCGCGCGCACCAACACAGTGGCGCTACTTCAGCGCGGAGATCCCGGCC 480
QY 160 IleSerThrArgGlyGlyGlyThrTrpLysGlyAlaPheGlnGlnAlaAlaGlnIle 179
DB 481 ATCACCTACCGCGTGGTGGCACTATACCAAGGGCGCTTCAGCAAGCGCGCAAAATC 540
QY 180 LeuLeuHisAlaArgGluLeuSerThrLysValValPheLeuIleThrAspGlyTrpSer 199
DB 541 CTTGCTCACTTAGAGAAAATCCACCAAAAGTCATATTTCTCATCACCGAGCGCTATTC 600
QY 200 AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe 219
DB 601 ATGGCGGAGACCCAGACCTATTCAGCATTCCTTCGGATTCGGAGTGGAGATCTTC 660
QY 220 ThrPheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLys 239
DB 661 ACCTTCGGGATTTGGCAGGGGAATATCCGGGAATGCAATGCAATGCTTCACCCCGAAG 720
QY 240 GluGluHisCysTrpLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAla 259
DB 721 GAAGAACATTTTACCTGCTCCACAGTTCCTTGAAGAAATTTGAGGCTTTAGCTCGCAGGGG 780
QY 260 LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTrp 279
DB 781 TTGCATGAGATCTTACCTTCTGGAGTTCATTCAGAGATATGGCCCGCTCTCTTAT 840
QY 280 LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHis 299
DB 841 CTCTGTGAGGCTGGGAAAGACTGCTGTGACAGAAATGCCAGCTGCAAAATGTGGGACACAC 900
QY 300 ThrGlyHisPheGluCysIleCysGluLysGlyTrpGlyLysGlyLeuGlnTrpGlu 319
DB 901 ACCGGTCAATTTGAATTCATCTGTGAGAGCGGCTATTACGGGAAAGGTCTGCGAGCATGAG 960
QY 320 CysThrAlaCysProSerGlyThrTrpLysProGluGlySerProGlyLysIleSerSer 339
DB 961 TGCACAGCTTGGCCATCAGGACATATTAAGCCCGGAGCTTCTCCAGGAGGAATCAGACCC 1020
QY 340 CysIleProCysProAspGluAsnHisThrSerProGlySerThrSerProGluAsp 359
DB 1021 TGCATCCCATGTCTGAGCTAAGCCACACCTCCCGACCTTGGAGGACATCTCCCTCGAAGAC 1080
QY 360 CysValCysArgGluGlyTrpArgAlaSerGlyGlnThrCysGluLeuValHisCysPro 379

DB 1081 TGGTGTGCGGAGAGGATACCAAGATCTGGCCAGACTGTGAGGTGTCCACTGTCT 1140
QY 380 AlaLeuLeuProProGluAsnGlyTrpPheIleGlnAsnThrCysAsnAsnHisPheAsn 399
DB 1141 GCGCTGAAGCTCTGTAAATGTTTTTATATAAAACACACTTGCAGAAACCACTTCAAT 1200
QY 400 AlaAlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeu 419
DB 1201 GCGCGCTGTGGGTCCGATGTGCCCGGCTTTGACCTTTGTGGGAAGCAGCATTCATTTG 1260
QY 420 CysLeuProAsnGlyLeuTrpSerGlySerGluSerTrpCysArgValArgThrCysPro 439
DB 1261 TGTCAACCAATGTTTGTGGTCTGGCAGAGAAGCTTCTGCAGATGAGACGTGCCCC 1320
QY 440 HisLeuArgGlnProLysHisIleSerCysSerThrArgGluMetLeuTrpLys 459
DB 1321 CACCTCCGACAGCCCAACAGCGCCACATCAGCTGTCCACTGCGGAATGTCTACAC 1380
QY 460 ThrThrCysLeuValAlaCysAspGluGlyTrpArgLeuGluGlySerAspLysLeuThr 479
DB 1381 ACCCTGTGTTGGTTACTTGCATGAAGGATACAGATTAGAGGAGCAGCAGCTAGCTTACC 1440
QY 480 CysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSer 499
DB 1441 TGTCAAGGAATGCCAGTGGATGCCAGAGCCCGGTGTGTAGAACCCACTTGTGCC 1500
QY 500 ThrPheGlnMetProLysAspValIleIleSerProHisAsnCysGlyLysGlnProAla 519
DB 1501 ACCTTCCAGAGCCCAAGAGCGTCATCTTCTCCAGCTGCGGCAAGCAGCCTGCC 1560
QY 520 LysPheGlyThrIleCysTrpValSerCysArgGlnGlyPheIleLeuSerGlyValLys 539
DB 1561 AGGCTGGGATGACCTGTGAGTAAAGTGGCGGCGAGGATACATTTTATCGGGGTGAGA 1620
QY 540 GluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCys 559
DB 1621 GAA--GTGAGATGTCCTCATCTGGAGTGGAGTGGCCAAAGTTCAGACAGCTGTGTGC 1677
QY 560 LysAspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGlu 579
DB 1678 AAAGATGTGGAGCTCCCAAAATCAGCTGTCCAAATGACATTTGAGGCAAGACCTGGGAG 1737
QY 580 GlnGlnAspSerAlaValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGlu 599
DB 1738 CAGCAGAGCTCTGCTAATGTACCTGGCAAGTCCCAACAGCTAAAGACAACTCTGGTGA 1797
QY 600 LysValSerValHisValHisProAlaPheThrProTrpTrpLeuPheProIleGlyAsp 619
DB 1798 AAGGTGTAGTCCAGCTCCAGCGCTTTACCCCACTTACCTTCCCAATTTGGAGAC 1857
QY 620 ValAlaIleValTrpThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHis 639
DB 1858 GTGGCCATCACTTACAGCGCAACCGACTCATCCGTAACCAAGCCAGCTGCTCTTCTAC 1917
QY 640 IleIysValIleAspAlaGluProProValIleAspTrpCysArgSerProProVal 659
DB 1918 ATTAAAGTCAATGATGTGAACCGCTGTATAGATTGGTGGCCGATCTTCCACCTCCCAATC 1977
QY 660 GlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSer 679
DB 1978 CAGGTGTAGAGGAGGACCCCTGCDAGCTGGGATGGAGCTCAGTTCTCAGACACTCC 2037
QY 680 GlyAlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGlu 699
DB 2038 GGGCTGAATTTGTTTATTTACAGCAGTCACACAGAGCGGACATGTTTCTCTCATGGGAA 2097
QY 700 ThrIleValGlnTrpThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHis 719
DB 2098 ACCGTGTGGTGTGATCAGCAGCTGACCTCCAGCGCAACAGAGACTGTGAGCATCCAC 2157
QY 720 IleValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIle 739

Db	338	ATGTGGCTCGCTGGCCCTTTTGTGTCTGGGGTCTGGCGCTCTGTTTCGGGCTGGGGGACC	397
Qy	21	PheGlnGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro	40
Db	398	TTTCAGCAGATGTCCCGTGGCGCAATTTTCAGCTTCGGCTCTCCCGGAGACCGCGCC	457
Qy	41	GlyAlaProGlySerIleProAlaProProAlaProGlyAspGluAlaAlaGlySerArg	60
Db	458	GGGGCCCCGGGAGTATCCCGCGCGCGCCCTCTCTGGCGAGCAAGCGCGGGAGCAGA	517
Qy	61	ValGluArgGluGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu	80
Db	518	GTGGAGCGGCTGGGCCAGCGCTTCGGCGACCGGTGGCGCTCTCGGGAGCTCAGCGAG	577
Qy	81	ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg	100
Db	578	CGCCTGGAGCTTGTCTTCCTGGTGATGATTCGTCCAGCGTGGCGCAAGTCAACTTCGCG	637
Qy	101	SerGluLeuMetPheValArgIleLeuSerAspPheProValValProThrAlaThr	120
Db	638	AGCGAGCTCATGTCGTCCGAAGCTGTCTCCGACTTCGCCGTGGTCCCGCCAGCCAGG	697
Qy	121	ArgValAlaIleValThrPheSerSerIleValValProArgValAspIle	140
Db	698	CGCGTGGCATCTGTACCTCTCTGTCGAAGAACTACGTGGTGGCGCGCTCGATTACATC	757
Qy	141	SerThrArgAlaArgGlnHisLysCysAlaLeuLeuLeuGlnGluIleProAlaIle	160
Db	758	TCCACCGCGCGCGCGCCAGCAAGTGGCGGTCTCTCCAGAGATCCCTGCCATC	817
Qy	161	SerTyrArgGlyGlyGlyThrTyrThrIleValPheGlnGlnAlaAlaGlnIleLeu	180
Db	818	TCCTACCGAGGTGGCGGCACCTACACCAAGGGCGCTTCACGCAAGCGCGCAATTCCT	877
Qy	181	LeuHisAlaArgGluAsnSerThrIleValPheLeuIleThrAspGlyTyrSerAsn	200
Db	878	CTTCATGTCTAGAGAAACTCAACAAGTTGTATTCATCATCTGATGGATATCCCAT	937
Qy	201	GlyGlyAspProArgProIleAlaSerLeuArgAspSerGlyValGluIlePheThr	220
Db	938	GGGGGAGACCTTAGACCAATTCAGCGTCACCTGCGAGATTTCAGGAGTGAGATCTTCACT	997
Qy	221	PheGlyIleThrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGlu	240
Db	998	TTTGGCATATGCAAGGGAAACATTGAGAGCTGAATGACATGGCTTTCACCCCAAGGAG	1057
Qy	241	GluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAlaLeu	260
Db	1058	GAGCATGTACTGTACACAGTTTGAAGAAATTGAGCTTTAGCTCGCGGCAATG	1117
Qy	261	HisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrLeu	280
Db	1118	CATGAAGATCTACCTTCTGGGAGTTTATATCAAGATGATATGCTCCACTGCTCATATCTT	1177
Qy	281	CysAspGluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThr	300
Db	1178	TGTGATGAAGCAAGAGACTGCTGTGACCAATGGGAAGCTGCAAAATGTGGGACACACA	1237
Qy	301	GlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGluCys	320
Db	1238	GGCCATTTTGAAGTCATCTGTGAAAGGGTATTACGGGAAGGCTCGCAGTATGAATGC	1297
Qy	321	ThrAlaCysProSerGlyThrTyrLysProGluLysProGlyGlyIleSerSerCys	340
Db	1298	ACAGCTTCCCATCGGGACATACAAACCTGAAGGCTCACAGGAGGAATCAGCAGTTGC	1357
Qy	341	IleProCysProAspGluAsnIleThrSerProGlySerThrSerProLysAspCys	360
Db	1358	ATTCCATGTCTGATGAATAATCACACCTCTCCACTGGGAAGCACATCCCTCGAGACTGT	1417
Qy	361	ValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAla	380
Db	1418	GTCTGCAGAGAGGATACAGGGCATCTGGCCAGACCTGTGNACTTGTCCACTGCTCC	1477
Qy	381	LeuLysProProGluAsnGlyTyrPheIleGlnAsnThrCysAsnAsnHisPheAsnAla	400
Db	1478	CTGAAGCCTCCCGAAAAATGTTTACTTTATCCAAAAACACTTGCACAAACCACTTCAATGCA	1537
Qy	401	AlaCysGlyValArgCysHisProGlyPheAspLeuValGlySerSerIleLeuLeuCys	420
Db	1538	GCCTGTGGGTCGGATGTCACTCTGATTTGATCTTGTGGGAAGCAGCATCATCTTATGT	1597
Qy	421	LeuProAsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHis	440
Db	1598	CTACCAATGTTTGTGGTTCGGCTTCAGAGAGCTACTGCAGAGTAAGAACAATGCTCTCAT	1657
Qy	441	LeuArgGlnProLysHisGlyHisIleSerCysSerThrArgGluMetLeuTyrLysThr	460
Db	1658	CTCCGCCAGCCGAAACATGGCCACATCAGCTGTCTTACAAAGGAAATGTTATATAGACA	1717
Qy	461	ThrCysLeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCys	480
Db	1718	ACATGTTTGGTTCCTGTGTATGAGGTACAGCTAGAGAGCGTGTATTAAGCTTACTTGT	1777
Qy	481	GlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThr	500
Db	1778	CAAGGAAACAGCCAGTGGGATGGCCAGAAACCCCGGTGTGGAGCGCCACTGTTCACC	1837
Qy	501	PheGlnMetProLysAspValIleSerProHisAsnCysGlyLysGlnProAlaLys	520
Db	1838	TTTTCAGATGCCCAAGATGTCTATCATATCCCCCAACTGTGTGGCAAGCAGCCAGCCAA	1897
Qy	521	PheGlyThrIleCysTyrValSerCysArgGlnGlyPheIleLeuSerGlyValLysGlu	540
Db	1898	TTTGGAGCACTCTGTATGTAAATGTCGCCCAAGGGTTTCATTTATCTGGAGTCAAGAA	1957
Qy	541	MetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGlnAlaAlaValCysLys	560
Db	1958	ATGCTGAGATGTACCACTTCTGAAAAATGGAATGTGGAGTTTCAGGAGCTGTGTATAA	2017
Qy	561	AspValGluAlaProGlnIleAsnCysProLysAspIleGluAlaLysThrLeuGluGln	580
Db	2018	GACGTGGAGGCTCTCNAATCAACTGTCTTAAGACATAGAGGCTAAGACTCTGGAAACAG	2077
Qy	581	GlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAspAsnSerGlyGluLys	600
Db	2078	CAAGATTTGCGCAATGTACTCTGGCAGATTCCACAGCTTAAGACAACTCTGCTGAAGAG	2137
Qy	601	ValSerValHisValHisProAlaPheThrProProTyrLeuPheProIleGlyAspVal	620
Db	2138	GTGTCACTCCAGTTCATCCAGCTTTCACCCCACTTACCTTTTCCCAATGGAGATGT	2197
Qy	621	AlaIleValTyrThrAlaThrAspLeuSerGlyAsnGlnAlaSerCysIlePheHisIle	640
Db	2198	GCTATCGTATACACGGCAACTGACCTATCCGGCAACCCAGCGAGCTGCATTTTCCATATC	2257
Qy	641	LysValIleAspAlaGluProProValIleAspTyrCysArgSerProProValGln	660
Db	2258	AAAGTTATGTAGCAGAACCACTGTCTATAGACTGGTGCAGATCTCCACTCCCGTCCAG	2317
Qy	661	ValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPheSerAspAsnSerGly	680
Db	2318	GTCTCGGAAGGTACATGCGCAAGCTGGATGAGCCTCAGTTCTCAGACAACTCAGGG	2377
Qy	681	AlaGluLeuValIleThrArgSerHisThrGlnGlyAspLeuPheProGlnGlyGluThr	700
Db	2378	GCTGAATTTGTCATTTACAGAAAGTCATACAAAGGAGACTTTTCTCCCTCAGGGGAGACT	2437
Qy	701	IleValGlnTyrThrAlaThrAspProSerGlyAsnAsnArgThrCysAspIleHisIle	720
Db	2438	ATAGTACAGTATACAGCCACTGACCCCTCAGGCAATACAGGCATGTGTATCCATATT	2497
Qy	721	ValIleLysGlySerProCysGluIleProPheThrProValAsnGlyAspPheIleCys	740
Db	2498	GTCAATAAAGGTTCTCCCTGTGAATTCATTTCAACCTGTAAATGGGGATTTTATATGC	2557


```
Qy 1978 AsnPheThrPheArgAsnThrValThrTyThrCysLysGluGlyTyThrLeuAlaGly 1997
Db 722 AACTTTCACTTTTGGGAACACAGTTGCTTACATCAAGAGGGGTACACCTTCTGCG 781
Qy 1998 LeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnCysLeu 2017
Db 782 CTGACACATCTGATGACAGGCAACCGCAATGGAATTCAGTAACACCAAGTGCCTG 841
Qy 2018 AlaValSerCysAspGluProProIleValAlaPheHisAlaSerProGluThrAlaHisArg 2037
Db 842 GCTGTCTCTGTGACGAGCCCCCAATGTGGACCAAGCCTCTCCACAGACTGCTCACAGG 901
Qy 2038 LeuPheGlyAspIleAlaPheTyThrCysSerAspGlyTyThrSerLeuAlaAspAsnSer 2057
Db 902 CTCCTTGGAGACACCGCGTTTACTACTGTGCGGATGGTTACAGCTCGCTGATAAATTC 961
Qy 2058 GlnLeuLeuCysAsnAlaGlnGlyLysTrpValProGluGlyGlnAspMetProArg 2077
Db 962 CAGCTCATCTGCATGCCAGGGGAACCTGGTTCCCTCGCGGGCCAGGCTGTGCGCGC 1021
Qy 2078 CysIleAlaHisPheCysGluLysProProSerValSerTyThrSerIleLeuGluSerVal 2097
Db 1022 TGCATAGCTCACTCTGTGAAAAACCCCATCTGTTTCTTACAGCATCTTGGAAATCTGTG 1081
Qy 2098 SerLysAlaLysPheAlaGlySerValValSerPheLysCysMetGluGlyPheVal 2117
Db 1082 AGCAAGCAAAAGTTTGCAGCTGGCTCGGTAGTGAAGCTTCAAGTGTATGGAGGGTTTGTG 1141
Qy 2118 LeuAsnThrSerAlaLysIleGluCysMetArgGlyGlyGlnTrpAsnProSerProMet 2137
Db 1142 CTGAACACCTCAGCGAAGATTGAATGCTTGAGAGGTGGAGGTGAGGCGCTTCTCCCTC 1201
Qy 2138 SerIleGlnCysIleProValArgCysGlyGluProProSerIleMetAsnGlyTyAla 2157
Db 1202 TGGTCCAGTGATCCCGGTGCGTGGAGAGCGCTCCCAAGCATCGCAAAATGGCTACCG 1261
Qy 2158 SerGlySerAsnTyThrPheGlyAlaMetValAlaTyThrSerCysAsnLysGlyPheTy 2177
Db 1262 AGTGGGCAAACTACAGTTTGGGGCGGTGGTGGCTTACAGCTGCCCAAGGGATTCTAT 1321
Qy 2178 IleLysGlyGluLysLysSerThrCysGluAlaThrGlyGlnTrpSerProIlePro 2197
Db 1322 ATCAAGGGGGAGAGAGAGACAGTGTAGGCCACAGACAGTGGAGTAAACCCACGCGC 1381
Qy 2198 ThrCysHisProValSerCysGlyGluProProLysValGluAsnGlyPheLeuGluHis 2217
Db 1382 ACCTGCCATCTGTGTCCTGTAACGAGCCACTAAGGTTGAGAACCGGCTTCTCGAGCAC 1441
Qy 2218 ThrThrGlyArgIlePheGluSerGluValAlaArgTyThrGlnCysAsnProGlyTyThrLysSer 2237
Db 1442 ACCACTGGCAGGACCTTTGAGAGCGAGCAAGGTTTCCAGTGCACCCAGGCTATTAAGGCA 1501
Qy 2238 ValGlySerProValPheValCysGlnAlaAsnArgHisTrpHisSerGluSerProLeu 2257
Db 1502 GCCGAGTCTGTGTTTGTTCACAGCCATCCCACTGGCAGCAGCAGCCCTCTG 1561
Qy 2258 MetCysValProLeuAspCysGlyLysProProIleGlnAsnGlyPheMetLysGly 2277
Db 1562 TCTGCAACCCCTCTCAACTGTGGGAAACCCCTCCATTCAGAAATGGCTTTTGAAGCA 1621
Qy 2278 GluAsnPheGluValGlySerLysValGlnPhePheCysAsnGluGlyTyThrGluLeuVal 2297
Db 1622 GAAAGCTTTGAAGTAGGTGCAAGGTTTCAAGTTTGTCTGTAATGAGGGATATAGCTCGTT 1681
Qy 2298 GlyAspSerSerTrpThrCysGlnLysSerGlyLysTrpAsnLysLysSerAsnProLys 2317
Db 1682 GGTGATAATTTCTTGACTTGGCAGAAATCTGGCAAAATGGAGTAAAGCAAGCCCGAG 1741
Qy 2318 CysMetProAlaLysCysProGluProProLeuLeuGluAsnGlnLeuValLysGlu 2337
Db 1742 TGTGTCCCAACCAAGTGTGAGAGCTCTCTCTTAGAAAAACCAAGCTGTATTAAGGAA 1801
Qy 2338 LeuThrThrGluValGlyValThrPheSerCysLysGluGlyHisValLeuGlnGly 2357
```

RESULT 7

AK052699

LOCUS

DEFINITION

AK052699 2736 bp mRNA linear HTC 20-SEP-2003
Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D630026K16 product:polydomain protein, full insert
sequence.

```
Db 1802 TTAGCTTCCGAGGTAGGAGTGATGACCATTTCTCTGAAGAGGGGCAATGCTTTCGAAGGC 1861
Qy 2358 ProSerValLeuLysCysLeuProSerGlnGlnTrpAsnAspSerPheProValCysLys 2377
Db 1862 CCTCTGTCTCTGAAGTGTGTCATCCGGCAATGGAATGGTTCCTTCTATTATGTAAG 1921
Qy 2378 IleValLeuCysThrProProLeuLysSerPheGlyValProIleProSerSerAla 2397
Db 1922 ATGGTCTTGTGCTCCCTTCCCTTGAATTCCTTGGCGTCTCCCTGCGTCTTCCCGTGT 1981
Qy 2398 LeuHisPheGlySerThrValLysTyThrCysValGlyGlyPhePheLeuArgGlyAsn 2417
Db 1982 CTTTCAATTTGGCAGTACTGTCAAGTATCTGTGTGTGCGAGGGTTCCTTAAGAGGCAGT 2041
Qy 2418 SerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeuProGluCysValPro 2437
Db 2042 CCACCATCTCTGCCAGGCTGATAGCACTGGAGTTCTCCATTGCCCGAATGTCGTTCG 2101
Qy 2438 ValGluCysProGlnProGluGluLeuProAsnGlyIleIleAspValGlnGlyLeuAla 2457
Db 2102 GTAGAATGTCCCAACCTGAGGAGATCTCTCAACGGTATCATCCAGTACACAGGGCTTGC 2161
Qy 2458 TyrLeuSerThrAlaLeuTyThrCysLysProGlyPheGluLeuValGlyAsnThrThr 2477
Db 2162 TATCTCAGCACCACTCTACACCTGCAAGCGGCTTGTAGTTAGTGGGCAATGCTACC 2221
Qy 2478 ThrLeuCysGlyGluAsnGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGlu 2497
Db 2222 ACCCTCTGTGGGAAATGGCCAGTGGCTGGAGGAAACCAATGTGCAACCCATTGAA 2281
Qy 2498 CysLeuLysProLysGluIleLeuAsnGlyLysPheSerTyThrAspLeuHisTyGly 2517
Db 2282 TGCCAGAGCCCAAGAGATTTTAAATGGCCAAATCTCTTCCGTGAGCTTTCAGTAGGA 2341
Qy 2518 GlnThrValThrTyThrSerCysAsnArgGlyPheArgLeuGluGlyProSerAlaLeuThr 2537
Db 2342 CAACCATCACATCTTTGTGACCGGGCTTCCGGCTCGAAGGTCCTCCAAATCCCTGACC 2401
Qy 2538 CysLeuGluThrGlyAspTrpAspValAspAlaProSerCysAsnAlaIleHisCysAsp 2557
Db 2402 TGTTTAGAGACAGGTGACTGGGATATGGATCCCCCTCTTGTGATGCCATCCACTCGAT 2461
Qy 2558 SerProGlnProIleGluAsnGlyPheValGluGlyAlaAspTyThrTyGlyAlaIle 2577
Db 2462 GACCCACAGCCCATTTGAAATGGTTTCTAGAGGTGGGATTACAGATACGTCGTCATG 2521
Qy 2578 IleIleTyThrCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGlu 2597
Db 2522 ATCATCTATAGCTGTTTCCCTGGGTTTCAAGTGTGCTTGTGTCATGCCATGCAGACTGTGAA 2581
Qy 2598 GluSerGlyTyTrpSerSerIleProThrCysMetProIleAspCysGlyLeuProPro 2617
Db 2582 GAGTCGGGATGCTCAAGCTCCAGCCCACTGTGTATCCCATAGACTGCGGTCTCCCTCCT 2641
Qy 2618 HisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyThrPheGluGlnGlu 2637
Db 2642 CACATAGACTTTTGGTGAATGTACTAAGTACAGAGTGGCAGGACATTTTGTATCAAGAA 2701
Qy 2638 AspAspMetGluValProTyThrProHisProProTyThrHisLeuGlyAlaVal 2657
Db 2702 GATGACATGATGGAAGTCCCATATCTGGCT-----CACCCCTCAACATTTTGAAGCAACA 2755
Qy 2658 AlaLysThrTrpGluAsnThrLysGluSerProAlaThr 2670
Db 2756 GCTAAGGCTTGGAAAAATACAAAGGAGTGGCTGCTCTCT 2794
```

AK052699
 VERSION AK052699.1 GI:26342892
 KEYWORDS HTC; CAP trapper;
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 20499374
 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2736)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koye, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Soejabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.
 URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.
 Location/Qualifiers
 1. .2736
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:D630026K16"
 /db_xref="MGI:2422522"
 /db_xref="taxon:10090"
 /clone="D630026K16"
 /tissue_type="kidney"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="0 day neonate"
 193. .>2736
 /note="unnamed protein product; polydomain protein (MGI:1928849, GB|NM_022814, evidence: BLASTN, 100%, match=1346)
 putative"
 /codon_start=1
 /protein_id="BAC35103.1"
 /db_xref="GI:26342893"
 /translation="MMSRLAPCCNALALVSGWTNFPQVAPSLNFSFLPLPEASPCALGRLAVPASSEERAGSKVERLGRAPRSVRRLRELSGSLVFLVDESSVGVQNFLELKFVRKLSDPFWATVAIVTSFSSKNVAVDYISTSAHQHCKALLSREIFAITYRGGTYTKGAFQAQAILRHSRENSTKVIPLITDYSNGGDPRIAAISLRDFGVEITFTGWOGNIRELNDMASTPKREHCYLLHSPFEALARRALHEDLPSSGFIOEDMARCSYLCEAGKDCDDMASCKGCTHTGQPCICEKGYGKGLQHECTACPKPPENGFFIQPGISTCIPCPDVSHSPSPGSPEDCVREGVQSGTCEVHVCHCPALPKPPENGFFIQNTCKHFNAAACGVRCPDFDLVSGSIHLCPNGLWSGTESFCVRVTCPLHQPQKHGHSCTAEMSYNTLCLVTNENYRLSTGLTTCQNAQWDGPPRCVHERHCATFQPKPGVITSPSCQKQPARPGMTQSLQSGVILSGREVRCATSGKWSAKVTAVCKDVEAPQISCNDIRAKTGQDQSANVTQVPTAKNSGKSVHVHPAFTPPVLPFGIDVAVITATDSGNOACSTFYIKVLIDVPPVIDWCESPPPIQVKEKHPASNDPQPSNDSGAE LUTSSHTQCDMPFHGETVWVTATDPSGNRRCTDHIIVTKGSPCEVFPVNGDFICADSGAVNCSLCKGIDYFTEGSTKYYCAPEDGIWRPPPYSTEWPCDCAIKRFRANHGPKSFEMLYKTKTRCDMDLFRKFAAFETTLGNMVPSPCNDADDDIDCRLEDLT"

ORIGIN
 Alignment Scores:
 Pred. No.: 3,6e-306 Length: 2736
 Score: 3955.00 Matches: 710
 Percent Similarity: 90.08% Conservative: 53
 Best Local Similarity: 83.83% Mismatches: 82
 Query Match: 19.80% Indels: 2
 DB: 11 Gaps: 2
 US-09-977-053-4 (1-3571) x AK052699 (1-2736)
 QY 1 MetTTPProArgLeuAlaPheCysCysTTPGlyLeuAlaLeuValSerGlyTTPAlaThr 20
 DB 193 ATGTGGTGGCGCTGGCCCTTTTGTGCTGGCTCTGGCACTGGTGTGAGCTGGACCAAC 252
 QY 21 PheGlnGlnMetSerProSerArgAnPheSerPheArgLeuPheProGluThrAlaPro 40
 DB 253 TTTCCAGCCCGTGGCCCTTCGCTCACTTCCAGCTTCCGCTTGTTCGCCGAGGCTCTCCG 312
 QY 41 GlyAlaProGlySerIleProAlaProAlaProGly---AspGluAlaAlaGlySer 59
 DB 313 GGGGGCTCTGGCGAGACTGGCGGTACCTCCCGCTCCAGTGAGGAGGAGGAGGAGGAGC 372
 QY 60 ArgValGluArgLeuGlyGlnAlaPheArgArgArgValArgLeuArgGluLeuSer 79
 DB 373 AAGTGGAGCGCTGGCGCGCGGTTCGGAGCGCGTGGCGGAGCTGGCGGAGCTCAGC 432
 QY 80 GluArgLeuGluValPheLeuValAspAspSerSerValGlyGluValAlaAspPhe 99
 DB 433 GGCAGCTGGAGCTGCTCTTCTCTGGTGGAGAGTCTGCCAGCTGGCGGAGGAGGAGGAGG 492
 QY 100 ArgSerGluLeuMetPheValArgIleLeuSerAspPheProValValProThrAla 119
 DB 493 CTCACAGAGCTCAAGTTCGTGGCAGCTGCTCCGACTTCCCGGTGGTGTCCACGGCC 552
 QY 120 ThrArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyr 139

Qy 3010 ProValIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyValAlaAlaArg 3029
 Db 123 CCCATTCATCAACGGGAACCATCAACGCAATCTGATTTGGGATGTGGAAAGACGGTCCAG 182
 Qy 3030 IleGlnCysPheIleGlyPheLeuLeuGlyLeuSerGluIleThrCysGluAlaAsp 3049
 Db 183 ATTGAGTCTCTCAAGGGCTTCAGCTGCTTGAGACTTCTGAAATCACTGTGATGCCAAT 242
 Qy 3050 GlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMet 3069
 Db 243 GGCCAATGTTCTGAC--GTCCACTGTGTGAGCAGCTCAGTGGCGGCTCTCCCAACC 299
 Qy 3070 IleProAsnAlaPheIleSerGluThrSerTrpLysGluAsnValIleThrTyrSer 3089
 Db 300 ATACCCAAACGCAATTTGCTTTGAGGCGACCTTTCCGAGGACAAATGTGGTAACTTACAGC 359
 Qy 3090 CysArgSerGlyTyrValIleGlnGlySerAspLeuIleCysThrGluLysGlyVal 3109
 Db 360 TGCAGACTGGCTACACCATGCAAGTAGTTCAGATCTGATTTGTACGGAAAGCGGATA 419
 Qy 3110 TrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAla 3129
 Db 420 TGGAGCCAGCTTACCCAAAGGTGTAACCCCTGCTCTGTGGACCCCAACCACTGTAGCC 479
 Qy 3130 AsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLeu 3149
 Db 480 AATGCAAGTGGCAACAGGAGAGGCTCATACCTATGAAGCAAGTGAATCTCAAGTGTCTG 539
 Qy 3150 GluGlyTyrThrMetAspThrAspThrPheThrCysGlnLysAspGlyArgTrp 3169
 Db 540 GAAGGATATGATGATTCGGATACAGATACATTCACCTGTCACCAAGATGGCCATTTGG 599
 Qy 3170 PheProGluArgIleSerCysSerProLysCysProLeuProGluAsnIleThrHis 3189
 Db 600 GTCCCTGGAAGAAATCACCTGCACTCTAAATAATGCTGTGCCATCCCAACATGACACGC 659
 Qy 3190 IleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAlaGlu 3209
 Db 660 ATACGTTTTTACCGAGATGACTTCCAGGTGAACAGACAAATTTCTGTGTCATGTGCGAA 719
 Qy 3210 GlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGluPro 3229
 Db 720 GGGTTTACCCAGAGAGTGAATCGTTCACATGCCAGCCGACGGTACATGGGAGCCA 779
 Qy 3230 PropheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGluHis 3249
 Db 780 CCATTTCTGATGATCTCTGATCCAGTGTGTTGGGCACTCTGAAAGCCAGCGCAT 839
 Qy 3250 GlyPheValValIleSerLysTyrThrPheGluSerThrIleIleTyrGlnCysGluPro 3269
 Db 840 GGCCTCGTGGTGGCAATGAACACAGCTTTGGAAGCACCATTGTTTACCAAGTGTGACCCT 899
 Qy 3270 GlyTyrGluLeuGluGlyValAsnArgGluArgValCysGlnLeuAsnArgGlnTrpSerGly 3289
 Db 900 GGTACAAATTAGGGGGAACAGGGAACCAATCTGCCAGAGACAGACAGTGGAGTGA 959
 Qy 3290 GlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeuAsnGlyLys 3309
 Db 960 GAGGTGGCAGTGTGCAGAGAGAACAAGTGTGAGACTCCAGCTGAGTTTCCCAATGGGAAG 1019
 Qy 3310 AlaAspIleGluAsnArgThrGlyProAsnValValTyrSerCysAsnArgGlyTyr 3329
 Db 1020 GCTGTCTGGAATAACCAACATCTGAGCCAGCTTCTGTTTCTCTGTCTCAGAGAGGCTAC 1079
 Qy 3330 SerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSerHisProVal 3349
 Db 1080 ACCCTGGAGGGTCCCCGAGGCACTGCACTGCAATGGAACTGGAAATCACCTGACT 1139
 Qy 3350 ProLeuCysLysProAsnProCysProValProPheValIleProGluAsnAlaLeuLeu 3369
 Db 1140 CCCCTCTGCAACCAATCCATCCCTGCTCTTTGTTGATCTCTGAGAAGCCGCTCTT 1199
 Qy 3370 SerGluLysGluPheTyrValAspGlnAsnValSerIleLysCysArgGluGlyPheLeu 3389

Db 1200 TCTGAAAAGAGTTTATGTGCACCAAGATGTATCTATCAAGTGCAGGAGGCTTCTCTG 1259
 Qy 3390 LeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGlnThrSerAla 3409
 Db 1260 CTCAAGGCAATGATGTGCATCAGCTGCAGCCCTGACGAGACATGACGACCAATGCC 1319
 Qy 3410 LysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIleAlaArgGly 3429
 Db 1320 AGATGTGAAAATAATCTCTCTGTGCTCTCAAGTACGCTAGAGAAATGCAATTCCTCGAGA 1379
 Qy 3430 ValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyrMetLeuGlu 3449
 Db 1380 GTGATTAACCAAGTATGGGACATGATCACTACTCTCTGTACAGTGGCTACATGCTAGAA 1439
 Qy 3450 GlyPheLeuArgSerValCysLeuGluAsnGlyThrThrSerProProIleCysArg 3469
 Db 1440 GGTTCCTCGGAGTGTGCTAGAAAATGCAATGACACCATCTCTCTGTGTTCAGA 1499
 Qy 3470 AlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSer 3489
 Db 1500 GCTGTCTGTGCTGCTCCCATGTCAGAAATGAGGTGTCTGTCAACGTCCTCAATGCTCTCA 1559
 Qy 3490 CysProGluGlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeuProCysLeu 3509
 Db 1560 TGCCACAGAGCTGATGGAGCTCTCTGTGAGAGCCATATGATCATCTCTCTCTGTTG 1619
 Qy 3510 AsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGlyTyrThrGlySer 3529
 Db 1620 AATGCTGGCGCTGTGTGCTGCTTATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1679
 Qy 3530 ArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCysValArgPro 3549
 Db 1680 CGCTGTCTATACAGCTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1739
 Qy 3550 AsnArgCysHisCysLeuSerSerTrpThrGlyHisAsnCysSerArgLysArgArgThr 3569
 Db 1740 AACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799
 Qy 3570 Gly 3570
 Db 1800 GGG 1802

RESULT 9

AK035333 3242 bp mRNA linear HTC 18-SEP-2003
 Mus musculus adult male urinary bladder cDNA, RIKEN full-length
 enriched library, clone:9530018121 product:polydomain protein, full
 insert sequence.

ACCESSION AK035333
 VERSION AK035333.1 GI:26330611
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 3 99279253
 PUBMED 10349636

AUTHORS

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 11042159
 REFERENCE
 AUTHORS

546	Db	TTCCAGCCCGTGGCCCTTCGCTCAACTTTCAGCTTTCGCGCTGTGTTCCCGGAGCCCTCTCTCCG	605
41	Qy	GlyAlaProGlySerIleProAlaProProAlaProGly---AspGluAlaAaGlySer	59
606	Db	GGGCTCTGGCAGACTGGCGGTACCTCCCGGTCAGTGGAGGAGGCAGCAGGGAGC	665
60	Qy	ArgValGluArgLeuGlyGlnAlaPheArgValArgLeuLeuArgGluLeuSer	79
666	Db	AAAGTGGAGCGCTCTGGCGCGCGTTCGAGACGCGTGCAGCGACTCGCGGAGCTCAGC	725
80	Qy	GluArgLeuGluLeuValPheLeuValAspSerSerSerValGlyGluValAspPhe	99
726	Db	GGCAGCTTGGAGCTGCTCTCTCGTGGACGAGCTGCACGCTGGCGCAACCACTTC	785
100	Qy	ArgSerGluLeuMetPheValArgLeuLeuSerAspPheProValValProThrAla	119
786	Db	CTCAACGAGCTCAAGTTCTGGCGCAGCTGTGTCCGACTTCCCGCTGGTGTCCACGGCC	845
120	Qy	ThrArgValAlaIleValThrPheSerSerIysAsnTyrValValProArgValAspTyr	133
846	Db	ACGCGTGTGGCATTCTCTCATCCAAAGAACACGTGTGTGGCGCGCTGGATTC	905
140	Qy	IleSerThrArgArgAlaArgGlnHisIysCysAlaIleLeuLeuGlnGluIleProAla	159
906	Db	ATCTCCACCAAGCCGCGCGCAACAACAAGTGCCTACTCAGCCCGCGAGATCCCGGCC	965
160	Qy	IleSerTyrArgGlyGlyGlyThrTyrThrIysGlyAlaPheGlnGlnAlaAaGlnIle	179
966	Db	ATCACTACCGCGTGTGGCCCTTATACGAGGGCGCTTCCAGCAGAGCCCGCCAAATC	1025
180	Qy	LeuLeuHisAlaArgGluAsnSerThrIysValValPheLeuIleThrAspGlyTyrSer	199
1026	Db	CTTCGTCACTCTAGAGAAAACCTCCACCAAGTCATATTTCTCATCACCGACGGCTATTCC	1085
200	Qy	AsnGlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePhe	219
1086	Db	AAATGGCGGAGACCCGACCTATTTCGACGATCGCTTCGGGATTTCCGAGTGGAGATCTTC	1145
220	Qy	ThrPheGlyIleTyrGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProIys	239
1146	Db	ACGTTTCGGATTTGGCAGGGGAATATCGGGACTGATGACATGGCTTCCACCCCGAG	1205
240	Qy	GluGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgArgAla	259
1206	Db	GAAGAACATTGTTACCTGCTCCACAGTTTGAAGAAATTGAGGCTTTAGCTCGCAGGGCG	1265
260	Qy	LeuHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyr	279
1266	Db	TTGCATCAAGATCTACCTCTTCGGAGTATTATCCAGAGGATATGCGCCCGCTGCTCTTAT	1325
280	Qy	LeuCysAspGluGlyLysAspCysCysAspArgMetGlySerCysIysCysGlyThrHis	299
1326	Db	CTCTGTGAGGCTGGGAAGAAGCTGCTGTGACAGANTGGCCAGCTGCAAATGTGGGACAC	1385
300	Qy	ThrGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGlyLeuGlnTyrGlu	319
1386	Db	ACGGCTCAATTTGAATGCACTCTGTGAGAGGGCTATTACGGGAAGGCTTGCAGCATGAG	1445
320	Qy	CysThrAlaCysProSerGlyThrTyrIysProGluGlySerProGlyGlyIleSerSer	339
1446	Db	TGCACAGCTTGCCTCATCAGGACATATAAGCCGGAGCTTCTCCAGAGGAGAAATCAGCAC	1505
340	Qy	CysIleProCysProAspGluAsnHisThrSerProProGlySerThrSerProGluAsp	359
1506	Db	TGCATCCATGCTCTGACGTAAAGCCACCTTCCCACTGAAGCACTTCCCTTGAAGAC	1565
360	Qy	CysValCysArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeu-----	375
1566	Db	TGCGTGTGGCAGAGGGATACGAGAGATGCGCCAGACCTGTGAGGG-TAAGCACCCCTAC	1624
375	Qy	-----	375
1625	Db	TGCTCAACCAATGGGATCTAGAAAGCTGATCTTGATCTCTGTGGTGCAGAGTGTGGTGTG	1684

```

Qy 376 -----ValHisCysProAlaLeuLysProProGluAsn 386
Db 2765 TTTTGTCTCTCTCTGAATGTTAGTTGTCCACTGCTCCCTGAGCCTCTGAAAT 2824
Qy 387 GlyTyrPheLeuGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCysGlyValArgCys 406
Db 2825 GTTTTATACAAACACCTTCGAAACCACTTCATGCGCTGTGGGTCCGATGT 2884
Qy 407 HisProGlyPheAspLeuValGlySerSerLeuLeuLysLeuCysLeuProAsnGlyLeuTrp 426
Db 2885 CGCCCGGCTTTGACCTTGTGGAGCAGCATCCATTTGTGTGTCACACCAATGTTGTGG 2944
Qy 427 SerGlySerGluSerTyrCysArg 434
Db 2945 TCTGGACAGAAAGCTTCTGCAGA 2968

RESULT 11
BX390270 906 bp mRNA linear EST 08-MAY-2003
LOCUS BX390270 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1010YG21 5-PRIME, mRNA sequence.
ACCESSION BX390270
VERSION BX390270.1 GI:30463533
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: secre@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4991.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOBAG0252E09 CS02373 1&cluster=4991.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOBAG0252E09_CS02373_1.
FEATURES
Location/Qualifiers
1..906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1010YG21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores: 7,45e-125 Length: 906
Pred. No.: 1690.00 Matches: 294
Score: 98.0% Conservative: 1
Best Local Similarity: 97.6% Mismatches: 6
Query Match: 8.46% Indels: 0
DB: 13 Gaps: 0

US-09-977-053-4 (1-3571) x BX390270 (1-906)

Qy 243 CysTyrLeuHisSerPheGluPheGluAlaLeuAlaArgAlaLeuHisGlu 262
Db 2 TGTACCTGCTACAGATTTTGAAGAAATTTAGGCTTTAGCTCGCGGCAATGCAATGAA 61
Qy 263 AspLeuProSerGlySerPheLeuGlnAspMetValHisCysSerTyrLeuCysAsp 282

```

```

Db 62 GATCTACCTTCTGGGAGTTTATTTCAGATGATATGGTCCACTGCTCATATCTTTGTGAT 121
Qy 283 GluGlyLysAspCysCysAspArgMetGlySerCysLysCysGlyThrHisThrGlyHis 302
Db 122 GAAGGCAAGGAGCTGCTGTGACCGAATGGAGCTGCAAAATGTGGGACACACACAGGCCAT 181
Qy 303 PheGluCysLeuLeuGlyGlyTyrTyrGlyLysGlyLeuGlnTyrGluCysThrAla 322
Db 182 TTTGAGTGCATCTGTGAAGAAGGGGTATTACGGGAAAGGTCTGCAATGATGAATGCACAGCT 241
Qy 323 CysProSerGlyThrTyrLysProGluGlySerProGlyGlyLysSerCysLeuPro 342
Db 242 TGCCCATCGGGACATACAACTGAGCTCACCAGGAGGAATCAGCAGTTGCATTCGA 301
Qy 343 CysProAspGluAsnHisThrSerProGlySerThrSerProGluAspCysValCys 362
Db 302 TGTCTGTATGAAATACACACCTCTCCACTGGAAGCACATCCCTGGAAGACTGTGTCTGC 361
Qy 363 ArgGluGlyTyrArgAlaSerGlyGlnThrCysGluLeuValHisCysProAlaLeuLys 382
Db 362 AGAGAGGGGATACAGGGCATCTGGCCAGACCTTGTGAACCTTGTCCACTGCCCCCTGAG 421
Qy 383 ProProGluAsnGlyTyrPheLeuGlnAsnThrCysAsnAsnHisPheAsnAlaAlaCys 402
Db 422 CTCTCCGAAATATGGTTACTTTATCCAAACACTTGCACACCACTTCAATGCAGCCTGT 481
Qy 403 GlyValArgCysHisProGlyPheAspLeuValGlySerSerLeuLeuCysLeuPro 422
Db 482 GGGGTCCGATGTCACCTGGATTGATCTTGTGGGAAGCAGCATCATCTTATGCTTACCC 541
Qy 423 AsnGlyLeuTrpSerGlySerGluSerTyrCysArgValArgThrCysProHisLeuArg 442
Db 542 AATGTTTGTGTGTCAGTTCCAGAGACTACTGCAGATGAAGAACATGCTCTCATCTCCGC 601
Qy 443 GlnProLysHisGlyHisLeuSerCysSerThrArgGluMetLeuTyrLysThrCys 462
Db 602 CAGCCGAAACATGGCCCATCATGAGCTGTTCTACAGGGAATGTTATATAGACACATGT 661
Qy 463 LeuValAlaCysAspGluGlyTyrArgLeuGluGlySerAspLysLeuThrCysGlnGly 482
Db 662 TTGGTTGCTGTGATGAAGGGTACAGACTAGAAGCAGTCAGTAAGCTTACTTGTCTAAGGA 721
Qy 483 AsnSerGlnTrpAspGlyProGluProArgCysValGluArgHisCysSerThrPheGln 502
Db 722 AACAGCCAGTGGGATGGGCCAGAACCCCGTGTGTGGAGCCACTGTTCCACCTTTCAG 781
Qy 503 MetProLysAspValLeuLeuSerProHisAsnCysGlyLysGlnProAlaLysPheGly 522
Db 782 ATGCCCAAGATGTCATCATATTTCCCACTGTGGCAGCAGCAGCCAAATTTGGG 841
Qy 523 ThrLeuCysTyrValSerCysArgGlnGlyPheLeuSerGlyValLysGluMetLeu 542
Db 842 ACGATCTGCTATGATGTTGCGCCAGGTTCCATTTATTTTGTAGTCAAGGATTTGCTG 901
Qy 543 Arg 543
Db 902 Arg 904

RESULT 12
BX439277 1201 bp mRNA linear EST 15-MAY-2003
LOCUS BX439277 Homo sapiens PLACENTA Homo sapiens cDNA clone CS08009YML1
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX439277
VERSION BX439277.1 GI:30773765
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.

```

TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4989.r For
more information about this cluster, see
<http://www.genoscope.cns.fr/>
<http://p1g.cluster.cgi?seq=CS0DE009A060P1&cluster=4989.r>. Contact :
Peng Liang Email : liang@lifetech.com URL :
<http://fulllength.invitrogen.com/Invitrogen/Corporation/1600>
Paradise Avenue Genoscope sequence ID : CS0DE009A060P1.

FEATURES

```

Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CS0DB009YM1"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSORT 6; lbt strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched, into
double-strand cDNA was digested with Not I and cloned, into
the Not I and EcoRV sites of the pCMVSORT 6 vector.
Library was not normalized."

```

ORIGIN

Alignment Scores:	
Pred. No.:	4, 25e-122
Score:	1658.50
Percent Similarity:	90.63%
Best Local Similarity:	90.08%
Query Match:	8.30%
Dg.	13
Gaps:	4
Indels:	15
Mismatches:	24
Conservative:	2
Matches:	327
Length:	1201

US-09-977-053-4 {1-3571} x BX439277 {1-1201}

3003	QY	LeuProCysArgCysSer----	ThrProValIleGluTyrGlyThrValAsnGlyThrAsp	3021
		:::		
45	DB	GTACCGGTCGGGAATCCCGGGATCCAGTAAATTGAATATGGAACCTGTCAATGGGACAGAT	104	
3022	QY	PheAspCysGlyValAlaAlaAArgIleGlnCysPheIysGlyPheIysLeuLeuGlyLeu	3041	
105	DB	TTTGACTGTGGAAAGGACGCCGGGATTCAGTGTCTCAAGAGCTTCAGCTCTCCTAGGACTT	164	
3042	QY	SerGluIleThrCysGluAlaAspGlyVGLnTrpSerSerGlyPheProHisCysGluHis	3061	
165	DB	TCTGAATACACCTGTGAAGCCGATGGCCAGTGGAGCTCTGGGTTCCTCCCACTGTGAACAC	224	
3062	QY	ThrSerCysGlySerLeuProMetIleProAsnAlaPheIleSerGluThrSerSerTrp	3081	
225	DB	ACTTCTTGTTGTTCTCTTCCAAATCATACCAATATCGTTTCATCATGAGAGACCACTCTTGG	284	
3082	QY	LysGluAsnValIleThrTyrSerCysArgSerGlyTyrValIleGlnGlySerSerAsp	3101	
285	DB	AAGGAAATGTGATAACTTACAGCTGCAGTCTGGATATGTCATACAGGCGATTCAGAT	344	
3102	QY	LeuIleCysThrGlnLysGlyValTrpSerGlnProTyrProValCysGluProLeuSer	3121	
345	DB	CTGATTGTACAGAGAAAGGGGTATGGAGCCAGCTTATCCAGTCTGTGAGCCCTTGTC	404	
3122	QY	CysGlySerProProSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyrGlu	3141	
405	DB	TGTGGGTCCCCACCGTCTGTTCGCAATGCAATGCAATGGCACTGGAGAGGCACACACTATGAA	464	
3142	QY	SerGluValIysLeuArgCysLeuGluGlyTyrThrMetAspThrAspThrAspThrPhe	3161	
465	DB	AGTGAGGTGAAATCTCAGATGTCCTGGAAAGGTTATACGATGGATACAGTACAGATACATTC	524	
3162	QY	ThrCysGlnLysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCys	3181	

FEATURES

http://image.nlm.gov
Plate: NDAM482 row: e column: 11
High quality sequence stop: 729.
Location/Qualifiers

```

source
1..918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30398026"
/tissue_type="White Matter"
/dev_stage="unknown"
/lab_host="DH108-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned [EcoRV site is destroyed upon cloning]. Average
insert size 1.42 Kb. Library was constructed by
(invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 5,618-120 Length: 918
Score: 1629.50 Matches: 295
Percent Similarity: 97.69% Conservative: 1
Best Local Similarity: 97.36% Mismatches: 6
Query Match: 8.16% Indels: 6
DB: 14 Gaps: 1
US-09-977-053-4 (1-3571) x CD515150 (1-918)

QY 3263 llelletyrGlnCysGluProGlytyrGluLeuGluGlyAsnArgGluArgValCysGln 3282
DB 2 ATTATTATCATGCTGAGCTGGCTATGAACTAGAGGGGACAGGGAAGCTGTCTGCCAG 61

QY 3283 GluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrPro 3302
DB 62 CAGAACACACAGCTGGAGGTGGCAATATGCAAAAGAGACACAGGTGTGAAACTCCA 121

QY 3303 LeuGluPheLeuAsnGlyLysAlaIleGluAsnArgThrThrGlyProAsnValVal 3322
DB 122 CTGTGAATTTCTCAATGGGAAGCTGACATTTGAAACAGGACGACTGGACCCACAGTGGTA 181

QY 3323 TyrSerCysAsnArgGlytyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsn 3342
DB 182 TATTCCTGCAACAGAGGCTACAGTCTTGAAGGGCCATCTGAGGCACACATGCACAGAAAT 241

QY 3343 GlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValProPheVal 3362
DB 242 CGAATCTGGAGCACCCAGCTCCCTCTCTGCAAAACCAATCCATGCGCTGTCTCTTTTGTG 301

QY 3363 IleProGluAsnAlaLeuLeuSerGluLysGluPheTyrValAspGlnAsnValSerIle 3382
DB 302 ATTCCGGAGATGCTCTGCTGTCTGAAAGGAGTTTATTTGATCAGAAATGTGTCCATC 361

QY 3383 LysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGlu 3402
DB 362 AAATGTAGGAAGGTTTCTGCTGCGAGGCCACCGCATCATTAACCTGCAACCCCGACGAG 421

QY 3403 ThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProAlaHisVal 3422
DB 422 ACGTGACACACAGACGCGCAAAATGTGAAATAATCTCATGTGGTCCACGCTCACGTA 481

QY 3423 GluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCys 3442
DB 482 GAAATGCAATTTGCTCGAGCGGTACATTTATCATATGAGACATGATCACTACTCATGT 541

QY 3443 TyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrp 3462
DB 542 TACAGTGGATACATGTTGGAGGTTTCTCTGAGGAGTGTGTTTGTAGAAATGGAACATGG 601

QY 3463 ThrSerProIleCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCys 3482
DB 602 ACATCACTCTCTATTGTCAGAGCTGTCTGTGATTTCCATGTGCAATGGGGGCACTCTGC 661

QY 3483 GlnArgProAsnAlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGluGluPro 3502
DB 662 CAACGCCCAATGCTTTGTTCTGTCCAGAGGGCTGTGATGGGGCGGCTCTGTGAGAGACA- 720

```

```

QY 3503 lleCysIleLeuProCysLeuAsnGlyGlyArgCysValAla-ProTyrGlnCysAspCys 3522
DB 721 ATCTGCATTTCTTCCCTGTCTGAACGAGGTCGTGTGTGGCCCTTACCAGTGTGACTG 780

QY 3522 sPro-ProGlyTrpThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuA 3542
DB 781 CCGGCTCTGGCTGGAGCGGGTCTCGCTGTCTACACAGCTGTTCCAGTCTCCCTGCTTAA 840

QY 3542 snGly-GlyLysCysValArgPro-AsnArgCysHisCysLeu---SerSerTrpThrGln 3560
DB 841 ATGTGGGGAATGTGTAGACCAAAACCGATGTCACTGGTCTCTTTCTTCTTGGACGGG 900

QY 3560 Y 3560
DB 901 G 901

RESULT 14
BX431419 Homo sapiens 881 bp mRNA linear EST 15-MAY-2003
LOCUS BX431419 Homo sapiens PLACENTA Homo sapiens cDNA clone CS00E009YM11
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION BX431419
VERSION BX431419.1 GI:30775023
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by life technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4989.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG0132D02_CS01164_1&cluster=4989.r.
Contact: Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BAG0132D02_CS01164_1.
Location/Qualifiers
FEATURES
source
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE009YM11"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Alignment Scores:
Pred. No.: 1,568-113 Length: 881
Score: 1549.00 Matches: 281
Percent Similarity: 96.56% Conservative: 0
Best Local Similarity: 96.56% Mismatches: 10
Query Match: 7.76% Indels: 4
DB: 13 Gaps: 0
US-09-977-053-4 (1-3571) x BX431419 (1-881)

QY 3267 CysGluProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGln 3286
DB 10 TGTGAGCTGGCTT-GAATAGA-GGGACAGGAGAGTGTCTGCCAGGAGACAGACAG 67

```

QY 3287 TrpSerGlyValAlaIleCysLysGluThrArgCysGluThrProLeuGluPheLeu 3306
 |||||
 Db 68 TGAGTGGAGGGTGCATATGCAAGAGACAGGTGTGAACCTCACCCTGATTTCTC 127
 |||||
 QY 3307 AsnGlyLysAlaAspIleGluAsnArgThrThrGlyProAsnValValTyrSerCysAsn 3326
 |||||
 Db 128 AATGGGAAGCTGCATTTGAAACAGGACGACGTGGACCCCAACGCTGATATTCCTGCAAC 187
 |||||
 QY 3327 ArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAsnGlyThrTrpSer 3346
 |||||
 Db 188 AGAGGCTACAGTCTTGAAGGGCCCATCTGAGGCACTGCACAGAAATGGAACCTGGAGC 247
 |||||
 QY 3347 HisProValProLeuCysLysProAsnProCysProValProPheValIleProGluAsn 3366
 |||||
 Db 248 CACCCAGTCCCTCTGCAACCAATCAATGCGCTGCTCTTTGTGATTCGAGAT 307
 |||||
 QY 3367 AlaLeuLeuSerGlyLysGluPheTyrValAspGlnAsnValSerLeuLysCysArgGlu 3386
 |||||
 Db 308 GCTCTGCTGCTGMAAGGAGTTTATGTTGATCAGATGTGTCCATCAATGTAGGGA 367
 |||||
 QY 3387 GlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAsnProAspGluThrTrpThrGln 3406
 |||||
 Db 368 GGTTCCTGCTGAGGGCCACGCGCATCATTTACCTGCAACCCCGACGAGGTGGACAG 427
 |||||
 QY 3407 ThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisValGluAsnAlaIle 3426
 |||||
 Db 428 ACAGGCCCAATGTGAAAATATCTATGTGTCCACGACTCATCGTAGAAATGCATTT 487
 |||||
 QY 3427 AlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCysTyrSerGlyTyr 3446
 |||||
 Db 488 GCTCGAGGCGGTACATTTATCAATATGAGACATGATCACCTACTCATGTTACAGTGATAC 547
 |||||
 QY 3447 MetLeuGluGlyPheLeuArgSerValCysLeuGluAsnGlyThrTrpThrSerProPro 3466
 |||||
 Db 548 ATGTTGGAGGGTTCTCGAGGAGTGTGTTGTTTGAATAATGGAACATGACCTCCT 607
 |||||
 QY 3467 IleCysArgAlaValCysArgPheProCysGlnAsnGlyGlyIleCysGlnArgProAsn 3486
 |||||
 Db 608 ATTTGACAGAGCTGTCTGTGATTTTCATGTCAGATGGGGGCATCTGCCAACCCCAAT 667
 |||||
 QY 3487 AlaCysSerCysProGluGlyTyrMetGlyArgLeuCysGluGluProIleCysIleLeu 3506
 |||||
 Db 668 GCTTGTTCCTGCTCCAGAGGGCTGATGGGCGCTCTGTGAAGAACCAATCTGATTCCT 727
 |||||
 QY 3507 ProCysLeuAsnGlyGlyArgCysValAlaProTyrGlnCysAspCysProProGlyTyr 3526
 |||||
 Db 728 CCTGTCTGAACGAGT-CGCTGTGTGGCCCTTACAG-TGTGACTGCCCGCTGGCTGG 785
 |||||
 QY 3527 ThrGlySerArgCysHisThrAlaValCysGlnSerProCysLeuAsnGlyGlyLysCys 3546
 |||||
 Db 786 ACGGGTCTCGCTGTATACAGCTGTTTGCACTTNCCTGCTTAATGGTGGAAATGT 845
 |||||
 QY 3547 ValArgProAsnArgCysHisCysLeuSerSer 3557
 |||||
 Db 846 GGTAGGACCAACGATGTCACTGCTTCTTCT 878
 |||||
 RESULT 15
 BX356254
 LOCUS
 DEFINITION BX356254 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1010Y21 5-PRIME, mRNA sequence.
 ACCESSION BX356254
 VERSION BX356254.1 GI:30380055
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4991.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0D1010AD11Q1&cluster=4991.r. Contact :
 Feng Liang Email: fliang@lifetech.com URL : Invitrogen Corporation 1600
 http://fulllength.invitrogen.com/ Faraday Avenue Genoscope sequence ID : CS0D1010AD11Q1.
 Location/Qualifiers
 FEATURES
 source
 1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1010Y21"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with NotI and cloned into the NotI and EcoRV
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores: Length: 1201
 Pred. No.: 5,64e-107 Matches: 295
 Score: 1470.50
 Percent Similarity: 93.99% Conservative: 2
 Best Local Similarity: 93.35% Mismatches: 18
 Query Match: 7.36% Indels: 5
 DB: 13 Gaps: 0
 US-09-977-053-4 (1-3571) x BX356254 (1-1201)
 QY 1 MetTrpProArgLeuAlaPheCysTyrGlyLeuAlaLeuValSerGlyTrpAlaThr 20
 Db 102 ATGTGGCTCGCTGGCTTTTGTGCTGGGCTCGCGCTGTTTCGGCTGGCGGACC 161
 QY 21 PheGlnMetSerProSerArgAsnPheSerPheArgLeuPheProGluThrAlaPro 40
 Db 162 TTTTCAGCAGATGTCCTCGCTGGCAATTTTCAGCTTTCGCTTTCCTCGGAGACCGCGCC 221
 QY 41 GlyValProGlySerIleProAlaProAlaProGlyAspGluAlaAlaGlySerArg 60
 Db 222 GGGGCCCCGGGAGTATCCCCGGCGCCCGCTCTCGGCGACGAGCGCGGGGAGCAGA 281
 QY 61 ValGluArgLeuGlyGlnAlaPheArgArgValArgLeuLeuArgGluLeuSerGlu 80
 Db 282 ATGGAGCGGCTGGCGCAGGCTTCGGCGACGCGTGGCTGCTGCGGAGCTCAACGAG 341
 QY 81 ArgLeuGluLeuValPheLeuValAspAspSerSerValGlyGluValAsnPheArg 100
 Db 342 CGCTCGAGCTTGTCTTCCTGGTGATGATTCGTCAGCGTGGCGGAGTCAACTTCGCG 401
 QY 101 SerGluLeuMetPheValArgLysLeuLeuSerAspPheProValValProThrAlaThr 120
 Db 402 AGCGAGCTCATGTCTGTCGCAAGCTGCTGTCCGACTTCCCGTGGTGGCCCGCGCCACG 461
 QY 121 ArgValAlaIleValThrPheSerSerLysAsnTyrValValProArgValAspTyrIle 140
 Db 462 CGCTGGCCATCGTGACCTTCTCGTCCAAAACCTACGCTGGTGGCGCGCTCGATTACATC 521
 QY 141 SerThrArgAlaArgGlnHisLysCysAlaLeuLeuGlnGluIleProAlaIle 160
 Db 522 TCCACCGCGCGCGGCCA-CACAAATGCGCGCTGCTCTCTCCAAAGAGATCCCTGCCATC 580
 QY 161 SerTyrArgGlyGlyThrTyrThrLysGlyValaPheGlnGlnAlaAlaGlnIleLeu 180
 Db 581 TCCTACCGAGGTGGCGCACCTACCAAGGGCGCTTCCAGCAAGCGCGCCAAATTTCT 640
 QY 181 LeuHisAlaArgGluAsnSerThrLysValValPheLeuIleThrAspGlyTyrSerAsn 200

```

641 CTTTCATGCTAGAGAAACTCAACAAAGTTGTAATTTCTCATCTAGTGATGATATTCCTCAT 700
QY -GlyGlyAspProArgProIleAlaAlaSerLeuArgAspSerGlyValGluIlePheTh 220
Db GGGGGGAGACCCCTAGACCAATTCAGAGCTCATCGGAGATTCAGGAGTGAGATCTTCAC 760
QY -PheGlyIleTrpGlnGlyAsnIleArgGluLeuAsnAspMetAlaSerThrProLysGly 240
Db TTTTGGCATATGCGCAAGGAGCAATTCGAGAGCTGAATGACATGGCTTCCACCCCAAGGA 820
QY uGluHisCysTyrLeuLeuHisSerPheGluGluPheGluAlaLeuAlaArgAlaLe 260
Db GGAGCAGCTGTTACCTGCTACACAGTTTGAAGAATTTGAGGCTTAGCTCGCGGGCATT 880
QY uHisGluAspLeuProSerGlySerPheIleGlnAspAspMetValHisCysSerTyrIle 280
Db GCATCAAGATCTACCTCTCGGAGTCTTATTCARATGATATGCTCCACTCATATCT 940
QY uCysAspGluGlyAspCysCysAspArgMetGlySerCysLysCysGlyThrHisTh 300
Db TTGTGATGAAGAAA-GACTGCTGTGACCGAATKGRAGSTG-GMAATGKGRMMACAC 998
QY rGlyHisPheGluCysIleCysGluLysGlyTyrTyrGlyLysGly 315
Db 999 AAGCAATTTTGATGCTC-TGTAAACACGGKATTAOCGGAAAGGC 1043

RESULT 16
BX356253/c 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX356253 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1010YG21 3-PRIME, mRNA sequence.
ACCESSION BX356253
VERSION BX356253.1 GI:30378038
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4991.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1010AD11NP1&cluster=4991.r. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0D1010AD11NP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1010YG21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 6,58e-102 Length: 1201
Score: 1407.50 Matches: 298
Percent Similarity: 89.18% Conservative: 7

```

```

Best Local Similarity: 87.13% Mismatches: 30
Query Match: 7.05% Indels: 14
DB: 13 Gaps: 4
US-09-977-053-4 (1-3571) x BX356253 (1-1201)
QY 610 ThrProTyrLeuPheProIleGlyAspValAlaIleValTyrThrAlaThrAspLeu 629
Db 1002 TQKCCCCCT---TCCTTTCCAAAT---GAGATGTGCTATCGWTACACGCMACCTACCTATC 949
QY 630 SerGlyAsnGlnAlaSerCysIlePheHisIleLysValIleAspAlaGluProProVal 649
Db 948 CCG---AACCAGCCWCGCGGCATTATCCATATTAAGGTTATGTT-GCAGAACACACCTGTC 893
QY 650 IleAspTyrCysArgSerProProValGlnValSerGluLysValHisAlaAlaSer 669
Db 892 WTAGATGTGTGAGATCTCCACCTCCGTCAGTTCTCGAGAAAGGTACATGCGCAAGC 833
QY 670 TrpAspGluProGlnPheSerAspAsnSerGlyAlaGluLeuValIleThrArgSerHis 689
Db 832 TGGAT-GAGCCTCAGTTCTCAGACAACTCAGGGGCTGAATT-GTCATTACAGAGTCAT 775
QY 690 ThrGlnGlyAspLeuPheProGlnGlyGluThrIleValGlnTyrThrAlaThrAspPro 709
Db 774 ACACAGGAGM-CTTTTCCCTCAGGGAG-ACATATGATCAGTATACAGCACTGACCCC 717
QY 710 SerGlyAsnAsnArgThrCysAspIleHisIleValIleLysGlySerProCysGluIle 729
Db 716 TCAGGCNATAACAGGACATGTGATATCCATATTTGTCATAAAGGTTCTCCCTGTGAATT 657
QY 730 ProPheThrProValAsnGlyAspPheIleCysThrProAspAsnThrGlyValAsnCys 749
Db 656 CCAATTCACACCTGTAAATGG-GATTTTATATGCACTCCAGATAATACTCGAGTCAACTGT 598
QY 750 ThrLeuThrCysLeuGluGlyTyrAspPheThrGluGlySerThrAspLysTyrCys 769
Db 597 ACATTAACCTGCTTGAGGGCTATGNTTTCAGAGAGGCTCTACTGACAGATATTATTGT 538
QY 770 AlaTyrGluAspGlyValTrpLysProThrTyrThrGluTrpProAspCysAlaLys 789
Db 537 GCTTATGAAGATGGGCTCTGGAAACCAACATATACCACTGAATGGCCAGACTGTGCCAA 478
QY 790 LysArgPheAlaAsnHisGlyPheLysSerPheGluMetPheTyrLysAlaAlaArgCys 809
Db 477 AAACGTTTTCAAACCAACGGGTTCAAGTCTTGAATGATGTTCTTACAAAGAGCTCGTTGT 418
QY 810 AspAspThrAspLeuMetLysLysPheSerGluAlaPheGluThrThrLeuGlyLysMet 829
Db 417 GATGACACAGATCTGATGAAGAAGTTNTCTGAAGCAATTTGAGAGTTCCTGGGRRAAAKG 358
QY 830 ValProSerPheCysSerAspAlaGluAspIleAspCysArgLeuGluGluAsnLeuThr 849
Db 357 GTCCCATCATTTGTAGTGATGACAGAGGACATTCAGCGCAGACTGGRGAGAGCTGRC 298
QY 850 LysLysTyrCysLeuGluTyrAspTyrGluAsnGlyPheAlaIleGlyProGly 869
Db 297 GAAAATATTCCTAGAAATATAATTAAGTCTATGAATGAAATGGCTTTCGAATTTGACACAGG 238
QY 870 GlyTyrPheAlaAlaAsnArgLeuAspTyrSerTyrAspAspPheLeuAspThrValGln 889
Db 237 GCTGGGGTGCAGCTAATAGGCTGGATTTCTTACGATGACTTCTCGGACACTGTGCTA 178
QY 890 --GluThrAlaThrSerIleGlyAsnAlaLysSerSerArgIleLysArgSerAlaProL 909
Db 177 TAGAAACACGCAACAGCATCGGCAATGCCAGGCGTCAACGGATTAAGAGAGGTCGCCCAT 118
QY 909 euSerAspTyrLysIleLysLeuIlePheAsnIleThrAlaSerValProLeuProAspG 929
Db 117 TATCTGACTATAAAATTAAGTTAAATTTTAAATCATCAGT----GTGCCATTTACCGATG 62
QY 929 LuArgAsnAspThrLeuGluTrpGluAsn-GlnGlnArgLeuLeuGlnThrLeuGluThr 948
Db 61 AAAGAAATGATACCTCTGATCGGAAATATNCAGAACGACTCTCTTCAGACATTTGGAAACT 2

```


Db 1386 CGTTGCCGCTGGAACTATCCGCTCCCTCG-----TGCATT-----GCAACCTGTGG 1433
 Qy 2147 ygluProProSerIleMetAsnGlyTyrAlaSerGlySerAsnTyr----- 2162
 Db 1434 AGGACGCTGAGCACTTGGGTGGTGTGATCTGAGCCCGCTCCAGGCTTCATACC 1493
 Qy 2163 -----SerPheGlyAlaMetValAl 2169
 Db 1494 CAACAACCTAGACTGCACCTGGAGGATCTCATATCCCATCGGCTATGGTGACATATCA 1553
 Qy 2169 atYr-----SerCysAsnLysGlyPheTyr-----IleLysGlyGluLysGly 2183
 Db 1554 GTTCTGAATTTTCTACCGAGCTAATCATGACTTCTTGAAATTCAAATGGACCTTA 1613
 Qy 2183 sSerThrCysGluAlaThrGlyGlnTrpSerSerPro-----IlePr 2197
 Db 1614 CCACACAGCCCACTGATGATGGACAAATTTAGCGGCACGGATCTCCCGCGGCTGCTGAG 1673
 Qy 2197 oThrCysHisProVal----- 2202
 Db 1674 CACAACGCATGAACCCCTCATCCACTTTTATAGTGACCAATTCGCMAAACCGGCAAGGATT 1733
 Qy 2203 -----SerCysGlyGluProProLysValGluAs 2212
 Db 1734 TAAACTTGTCTTACCAAGCCTATGAATTACAGACTGTCCAGATCCACCCCATTTTCAGA 1793
 Qy 2212 nGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValAlaTyrGlnCysAs 2232
 Db 1794 TGGGTACATGATCAACTCGGATACAGCGTGGGCAATCA---GTATCTTTCGAGTGTTA 1850
 Qy 2232 nProGlyTyrLysSerValGlySerProValPheValCysGln-----AlaAsnArgHi 2250
 Db 1851 TCTTGGGTACATCTTAATAGGCCATCTCTGCTCTGTCAGCATGGGATCAACAGAA 1910
 Qy 2250 sTrpHisSerGluSerProLeuMetCysValProLeuAspCysGlyLysProProIle 2270
 Db 1911 CTGGAACTACCTTTTCCAGATGTGATGCCCT-----TGTGGGTACAACTTAATTC 1964
 Qy 2270 eGluAsnGlyPheMetLysGlyGluAsnPheGluValGlySerLysValGlnPheCys 2290
 Db 1965 TCAGAACGCCACCACTACTCCTCCCTGGCTTCTCT----- 1998
 Qy 2290 sAsnGluGlyTyrGluLeuValGlyAspSerSerTrp-----Thr-CysGlnLysSerG 2308
 Db 1999 ----GATGAGTATCCGATCCGAGGACTGCATTTGGCTCATCAGCGTCCCTCCAGGCA 2054
 Qy 2308 lLysTrpAsnLysLysSerAsnProLysCysMetProAlaLysCysProGluProL 2328
 Db 2055 CGGAGTTTACA-----TCACTTCACTCTGTACAGACGGAAGCTGTCAACGATTACAT 2108
 Qy 2328 euLeuGluAsnGlnLeuValLeuLys-----GluLeuThrThrGluValGlyValValThrP 2347
 Db 2109 TGCTGTTTGGGACGCTCCGATCAGAACTCAACCCAGCTGGGAGTTCAGTGGCAACAC 2168
 Qy 2347 heSerCysLysGluGlyHisValLeuGlnGlyProSerValLeuLysCysLeuProSerG 2367
 Db 2169 --AGCCCTCGAAACGGCTATAGCTCCACCAACCAAGT-----CTGTCTC 2211
 Qy 2367 ln-----GlnTrpAsn-----AspSerPhePro-ValCysLys 2377
 Db 2212 AAGTTCCACAGCGACTTTTCAATGGAGCTTCTTTGTCTCAATTTCCAGCATTTTCAG 2271
 Qy 2378 lleValLeuCysThrProProProLeuLeuSerPheGlyValProIleProSerSerAla 2397
 Db 2272 CTCAGAAATGTCACCTCCCTCCCA-----GCGTTCCTCCACAGCAGAA 2313
 Qy 2398 Leu-----HisPheGlySerThrValLysTyrSerCysValGly 2410
 Db 2314 ATGCTTACTGAGGATGATGATTTTCGAGATAGAGATTTTGTGAGTACAGTGCACCC 2373
 Qy 2411 GlyPhePheLeuArgGlyAsnSerThrThrLeuCysGlnProAspGly-----ThrTrp 2428
 Db 2374 GGGTACACCTTGTGGGGACCGACATCTCTGACTTGCAGACTCAGTTCCAGTGTGAGTTT 2433

Qy 2429 SerSerProLeuProGluCysValProValGluCysProGlnProGlu-----Gluile 2446
 Db 2434 GAGGTTTCTCTCCCAACATGT---GAAGCACAATCCCAAGCAAAATGAAGTCCGAGCTGGA 2490
 Qy 2447 ProAsnGlyIleIle-AspValGlnGlyLeuAlaTyrLeu-----SerThrAl 2462
 Db 2491 TCATCGGAGTCAATTTCTCAGTCAGGATATCGGGTATCGGGTAATATTTAACTCCCGAGCTGC 2550
 Qy 2462 aleuTyrThrCysLysProGlyPheGluLeuValGlyAsnThrThrThrLeu----- 2479
 Db 2551 TCTTGGAGTATTAAAG-----TGAACCAAACTACAAACATTTACATCTTT 2595
 Qy 2480 -----CysGlyGluAsnGlyHisTrp-----LeuGl 2488
 Db 2596 GTGGACACATTTCAAAGTGAAGAGCAGTTTGTATGCATGGAAGTGTTCATGGTTCTTCT 2655
 Qy 2488 yGlyLysProThrCys-----LysAlaIleGluCysLeuLysProLysGluIleLeuAs 2506
 Db 2656 GGGCAAGTCTCTCTCTAGTAGTCTTAAGTGGGAATCATCTGAACAAATCAATTTTACA 2715
 Qy 2506 n---GlyLysPheSerTyrThrAspLeuHisTyrGlyGlnThrValThrTyrSerCysAs 2525
 Db 2716 AGCAGAGTAAATCAGTTAT-----ATCTCCGCTGTCC-ACTGACCATGCCACAGTAA 2768
 Qy 2525 nArgGlyPheArgLeuGluGlyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAs 2545
 Db 2769 GAAGGATTTCAAGATTTCGC----- 2787
 Qy 2545 pValAspAlaProSerCysAsnAlaIleHisCysAspSerProGlnProIleGluAsnGl 2565
 Db 2788 -TATGCAGACCTTACTGAGTTTGCACCCAC-----CCCTGAAGATGG 2831
 Qy 2565 yPheVal-----GluGlyAlaAspTyrSerTyrGlyAlaIleIleTyrse 2581
 Db 2832 GGGTATTCTAAACAGGACTGCAGGAGCG-----GTTGGAAGCAAAATGCAATTATTT 2882
 Qy 2581 rCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer----- 2599
 Db 2883 TTGCAAGCTTGATACCAAGTGTGGCCACAGCAATGCACTGTAGACGAACCCACT 2942
 Qy 2600 -----GlyTrpSerSerSerIleProThrCysMetProIleAspCysGlyLeuPr 2616
 Db 2943 TGGCATGTACCAGTGGGACTCCTCAGCCACTCTCCAGCGCTGTCTCTGTGAATCCC 3002
 Qy 2616 oProHisIleAspPheGlyAspCysThrLysLeuLysAspAspGlnGlyTyrPheGluGl 2636
 Db 3003 A----- 3003
 Qy 2636 nGluAspAspMetMetGluValProTyrValThrProHisProTyrHisLeuGlyAl 2656
 Db 3003 ----- 3003
 Qy 2656 aValAlaLysThrTrpGluAsnThrLysGluSerProAlaThrHisSer-----Se 2673
 Db 3004 -----GAATCCCCAGGAAACGGTTTCATTTTACCGGAA 3035
 Qy 2673 rAsnPheLeuTyrGlyThrMetValSerTyrThrCysAsnProGlyTyrGluLeuGlu 2693
 Db 3036 CGAGTTTCACTTTGGACAGTAAAGTGTCTATGAATGTATGAGGGCTTCAAGCTGAATC 3095
 Qy 2693 yAsn-----ProValLeuIleCysGlnGluAspGlyThrTrpAsn-----GlySerAl 2709
 Db 3096 CAGCCAGCAAGCAACAGCGGTGTTCAGAGATGGGTGTGGAGTAACAGGGGAAGCC 3155
 Qy 2709 aProSerCysIleSerIleGluCysAspLeuProThrAla-----ProGluAsnGlyPh 2727
 Db 3156 GCCCATGTGTAAGCGGTGCTTGGCCCGCAGCATTTGAAGTCTCAGCTCTCAGAACATGCTAT 3215
 Qy 2727 eleuArgPheThrGluThrSerMet-----GlySerAlaValGlnTyrSerCysLy 2744
 Db 3216 CTGGAGGCTGGTTTCAGGATCTTGAATAGTACGGTGTCTCAAGTATTGCTGAGCTGCAG 3275

2744 sProGlyHisIleLeuAlaGlySerAspLeuArgLeuCysLeuGluAsnArgLysTrpSe 2764
3276 TCCTGGTTACTACTTAGAAGGCTGGAGGCTCTCGGGTGCAGGCCAATGGGACGTGAA 3335
2764 r-----GlyAlaSerProArgCysGluAlaIleSerCysLysLysProAsnProValMe 2782
3336 CATAGGAGATGAGAGGCGCAAGCTGTGAGCTGCGCACTGCGGTTCCCGACAGCCGATTTGT 3395
2782 tAnGlySerIleLysGlySerAsnTyrThrTyrLeuSerThrLeuTyrTyrGluCysAs 2802
3396 GAACGGTCACATTAGTGGAGATGGCTTCAGTTACAGACAGACCGGTGGTTTACCATGTCAA 3455
2802 pProGlyTyrValLeuAsnGlyThrGluArgArgThrCysGlnAspAspLysAsnTrpAs 2822
3456 TCCTGGTTTCGGCTTGTGGAACTTCGGTAGGATATGCTCTCAAGACACCAAGTGTGTC 3515
2822 pGluAspGluProIleCysIleProValAspCysSerSerProProValSerAlaAsnGln 2842
3516 TGGACAACGGCTGTCTGTGTCGCCATCATGTGGTCACTCGAATCTGGAACCTGCCACGG 3575
2842 yGlnValArgGlyAspGluTyrThrPheGlnLysGluIleGluTyrThrCysAsnGluGln 2862
3576 ATTCACTAATGGCAGTGAGTTCAACTGAATGATGTCTGCAATTCACCTGCAACACGGG 3635
2862 yPheLeuLeuGluGlyAlaArgSerArgValCysLeuAlaAsnGlySerTrpSerGlyAl 2882
3636 CTATTTGCTGCGGGGCTGTCTGAGCCAGTGTGCGAGCAACGGCCAGTGGAGTAGCCC 3695
2882 aThrProAspCysValProValArgCysAlaThrProProGlnLeuAlaAsnGlyValTh 2902
3696 TCTGCCCAAGCTGCGAGTGTGAAGTCTCTGATCCAGGCTTTGTGGAAATGCCATTCG 3755
2902 rGluGly-----LeuAspTyrGlyPheMetLysGluValThrPh 2915
3756 TCACGGGCAACAGAACTTCCTGAGAGTTTGTGATGGAAATG-----AGTATCTGTGA 3809
2915 eHisCysHisGlyGlyTyrIleLeuHisGlyAlaProLysLeuThrCysGlnSerAspGln 2935
3810 CCATTGCAAGAGGATTTTACTTGTGGGATCTTCAGCCTTGACCTGTATGCAAAATGG 3869
2935 yAsnTrpAspAlaGluIleProLeuCysLysProValAsnCysGlyProProGluAspLe 2955
3870 CTTATGGGACCGATCCCTGCGCAAGTGTGTGGCTATATCTGTGTGACACCCAGGGGTCCC 3929
2955 uAlaHisGlyPheProAsnGlyPheSerPheIleHisGlyGlyHisIleGlnTyrGlnCy 2975
3930 TGCCAAACGCGTCTCTCAGTGGAGAGTGTTTACTATGGGCGCGTGTGTCACACTCTCTG 3989
2975 sPheProGlyTyrLysLeuHisGlyAsnSerArgArgCysLeuSerAsnGlySerTr 2995
3990 CAGAGGAGCGAGAGCTCATAGGCAACGACAGAGGTGTGCGAGAGACAGTCACTG 4049
2995 pSerGlySerSerProSerCysLeu-----ProCysArgCysSerThrPro---Va 3011
4050 GAGCGGGGCACTGCCCACTGCACAGAAATATCTCGGATTCGTGTGTGATCCGGGAC 4109
3011 lIleGluTyrGlyThrValAsnGlyThrAspPheAspCysGlyLysAlaAlaArgIleGln 3031
4110 CCCAGACATGGGTCTCGGCTGGTGATGACTTTAAGACAAAGAGTCTTCTCGCTTCTC 4169
3031 nCysPheLysGlyPheLysLeuLeuGlyLeuSerGluIleThrCysGluAlaAspGlyGln 3051
4170 CTGTGAATGGGCGACCACTGAGGGGCTCCCTGGAACGCACTGTGTGCTCAATGGGTC 4229
3051 nTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuProMetIlePr 3071
4230 ATGGTCAGGACTCGACCGGCTGTGTGAGGCGCGGTCTCTGTGGCAACCTCGGCACACCCAC 4289
3071 oAsnAlaPheIle-----SerGluThrSerSerTrpLysGluAsnValIleThrTrpSe 3089
4290 CAACGGAATATGTTCAGTAGTAGGCAATTCGTTCCTCAGCTCGGTATC-----TATGC 4346
3089 rCysArgSerGlyTyrValIleGlnGlySerSerAspLeuIleCysThrGluLysGlyVa 3109

4347 CTGCTGGGAAGGCTTACAAGACCTCAGGCTCATGACAGCGCATTCACAGCCAAATGGGAC 4406
3109 lTrpSerGlnProTyrProValCysGluProLeuSerCysGlySerProProSerValAl 3129
4407 CTGACAGCAGCTGCTCCGCACTGCAATATTAAGTTGTGGGATCCAGGCACATAGC 4466
3129 aAsnAlaValAlaThrGlyGluAlaHisThrTyrGluSerGluValLysLeuArgCysLe 3149
4467 AAATGGCATCCAGTTTGGACCGACTTCACCTTCAACAGACTGTGAGCTATCAGTGTAA 4526
3149 uGluGlyTyrThrMetAspThrAspThrAsp---ThrPheThrCysGlnLysAspGlyYar 3168
4527 CCCAGGCTATGTATGCAAGCAGTCACTCCGCCACTATTTCGCTGTACAAAGACGGCAG 4586
3168 gTrpPheProGluArgIleSerCysSerProLysLysCysProLeuProGluAsnIleTh 3188
4587 GTGAATCCGAGCAAACTGTCTGCAAGCGCTGTGTCTTCAGCGCGCCGCGTGA 4646
3188 rHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerValSerCysAl 3208
4647 GAATGGACAGTGGAGGAAGTGAATTCCTGCGGCTCCAGCATAAAGTTACAGTGCAT 4706
3208 aGluGlyTyrThrPheGluGlyValAsnIleSerValCysGlnLeuAspGlyThrTrpGln 3228
4707 GGACGGTTACCAGCTCTCTCACTCCGCCACTCTCTCTGTGAAGTCCGCGGTGTGAA 4766
3228 uProProPheSerAspGluSerCysSerProValSerCysGlyLysProGluSerProGln 3248
4767 AGGAGAGATC-----CCCAAGTGTCTCCCTGTCTTCGCGAGACCCCTGGCATCCCCGC 4820
3248 uHisGlyPheValValGlySerLysTyrThrPheGluSerThrIleIleTyrGlnCysGln 3268
4821 AGAAGGCGCACTAGTGGGAAAGTTTCACTATAGTCCGAGTCTTCTTCCAGTGCAA 4880
3268 uProGlyTyrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnArgGlnTrpSe 3288
4881 ATCTCCATTTNN 4940
3288 rGlyGlyValala-----NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3292
4941 NNN 5000
3292 -----NN 3292
5001 NNN 5060
3292 -----NN 3292
5061 NNN 5120
3293 -----NN 3301
5121 NNN 5174
3301 rProLeuGluPheLeuAsnGlyLysAlaAsp-----IleGluAsnArgThrGln 3318
5175 CCCG-----GCACAGCGGATGTGAGAGCATCATCTTCTTACTTTCGG 5219
3318 yProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSerGluAlaHi 3338
5220 CTACACCTTAGTGTACACCTGCCATCCAGGCTTTTCTCGCAGCGGATCTGAGCACAG 5279
3338 s---CysThrGluAsnGlyThrTrpSerHisProValProLeuCysLys----- 3353
5280 AACATGTAAAGCAGACATGAATGAAGAGTCCGCTGTGTGTGTAAAGTAAAGAGT 5339
3354 -----ProAsnProCysProValProPheVal 3362
5340 GAGAGAGTTAATGAACAGATTACTAAACTCCAGTTCCTTCCAGATGTCTTTTTCGTC 5397

RESULT 18
AY407075

LOCUS AY407075 5901 bp DNA linear GSS 15-DEC-2003
DEFINITION Mus musculus CSMD1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY407075
VERSION AY407075.1 GI:39763046
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 5901)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5901)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES location/Qualifiers
source 1..5901
gene /organism="Mus musculus"
ORIGIN /mol_type="genomic DNA"
Alignment Scores: /db_xref="taxon:10090"
Pred. No.: 5,83e-99 Length: 5901
Score: 1386.50 Matches: 527
Percent Similarity: 34.80% Conservatives: 231
Best Local Similarity: 24.20% Mismatches: 751
Query Match: 6.94% Indels: 675
DB: 29 Gaps: 91
US-09-977-053-4 (1-3571) x AY407075 (1-5901)
Qy 1687 ArgIleSerCysGlyValProProLeuGluAanglyPheHisSerAlaAspPhe 1706
Db 190 CGGGAAGCTGCTTTGACCTCGGGAACATATGAATGGACAGGATTGGAACTGACTTT 249
Qy 1707 TyrAlaGlySerThrValThrTrpGlnCysAsnAnglyTyrtYrLeuLeuGlyAspSer 1726
Db 250 AAACCTGGCTCTACAGTTACCTATCAGTGTGATCTGTTGATACAGATTGTGATCCCTCG 309
Qy 1727 ArgMetPheCysThrAspAsnGlySerTrpAsnGlyValSerProSerCysLeuAspVal 1746
Db 310 TCCATTGAGTGTGT-GACAGGGGCTGATGGAGCGGTC----- 347
Qy 1747 AspGluCysAlaValGlySerAspCysSerGluHisAlaSerCysLeuAsnValAspGly 1766
Db 347 ----- 347
Qy 1767 SerTyrlleCysSerCysValProProTyrtYrThrGlyAspGlyLysAsnCysAlaGluPro 1786
Db 348 -----CTGGACCTCTGTGGAGCCAGTACACGGGC-----TCGGAGGGG 386
Qy 1787 IleLysCysLysAlaProGlyAsnProGluAanglyHisSerSerGlyGluIleTyrtYr 1806
Db 387 GTA---GTTTCTCACCACAACTACCTCATAC-----TACACA 422

Qy 1807 ValGlyAlaAlaValThrPheSerCysGlnGluGlyTyrtYrGlnLeuMetGlyValThrLys 1826
Db 423 GCTGGGCGAGATGTCATCTATTCC----- 446
Qy 1827 IleThr-CysLeuGluSerGlyGlyTyrAsnHisLeuIleProTyrtYrCysLysAla----- 1844
Db 447 ATCAGGGTGGCCCAAGGATTCGTGTATTTGGACAGTTTGCCTATTTCAGATGCACTG 506
Qy 1845 -----ValSerCysGlyLysProAlaIleProGluAangly-CysIleGluGlu 1861
Db 507 AATGACTTGGCAGAGTTGTTGATGGAACACATCCCGAGCCAGCGCTGT---GAAACAC 563
Qy 1861 euAlaPheThrPheGlySerLysValThrTyrtYrArgCysAsnLysGlyTyrtYrThrLeuAlaG 1881
Db 564 TCCCACTGGTACATCCCAATCAGATTCTGCTCGA-----TTGACGG 605
Qy 1881 LysPylsGlnSerSerCysLeuAlaAsnSer-SerTrpSerHisSerProProValCysG 1901
Db 606 CCAAAAGCGGAGCTTCTGCACGGGGTTTCCACTTCCTACCAAGCTGTTCCGGCACCA 665
Qy 1901 luProValLysCysSerSerProGluAsnIleAsnAsnGlyLysTyrlleLeuSerGlyL 1921
Db 666 GTGATACACAGTGCAGCTCTGTCCCGAGCCAGATATGGAGAGGATTTGTTCTCAGT 725
Qy 1921 euThrTyrtYrLeuSerThrAlaSerTyrtYrSerCysAspThrGlyTyrtYrSerLeuGlnGlyPro 1941
Db 726 TCTCCGCGAGCTCCATCGTCCGCTTCCAGTGTAAACCCAGGTACCTGCTGCAAGTTCCA 785
Qy 1941 erIleIleGluCys-----ThrAlaSerGlyIleTrpAspArgAlaProPro 1957
Db 786 CGGCATCGCTGCCAGTCTGTGCCAATGCTTTGGCTCAGTGGATGATGACACCATCCAA 845
Qy 1957 laCysHisLeuValPheCysGlyGluProProAlaIleLysAspAlaValIleThrGlyA 1977
Db 846 GCTGT-----GTAGTTCCATGCACTG 866
Qy 1977 snAsnPheThrPheArgAsnThrValThrTyrtYrThrCysLysGluGlyTyrtYrThrLeuAlaG 1997
Db 867 GCACCTTCACTCAGAGAGA-----G 887
Qy 1997 lylLeuAspThrIleGluCysLeuAlaAspGlyLysTrpSerArgSerAspGlnGlnCysL 2017
Db 888 GGAC-GATTTGTCCTCCAGGCTATCTCTGAGCCCTATGGGACACCACTGAA----- 936
Qy 2017 euAlaValSerCysAspGluProProIleValAspHisAlaSerProGluThrAlaHisA 2037
Db 937 -----CTGTGTGTGGAA-----GATCATAGT----- 957
Qy 2037 rgLeuPheGlyAspIleAlaPheTyrtYrCysSerAspGlyTyrtYrSerLeuAlaAspAsn 2057
Db 958 -----ATCGAGGGCTCGGGGATCCAGATCCAA 985
Qy 2057 erGlnLeuLeuCysAsnAlaGlnGlyLysTrpValPro-----Pro-GluGlyGlnAsp 2074
Db 986 GTGATCAGCTTTGCCACAGACGAGCACTGGGACTCCCTGGAGATCCATGATGAGGAGAC 1045
Qy 2075 Met-----ProArgCysIleAlaHisPhe----- 2082
Db 1046 ATGACGGCCCGGAGA---CTGGGCGACTTCTCAGNNNNNNNNNNNNNNNNNNNNNNNN 1102
Qy 2082 ----- 2082
Db 1103 NNN 1162
Qy 2083 -----CysGluLysPro-----ProSer 2088
Db 1163 NNN 1222
Qy 2089 valSerTyrtYrIleLeuGluSerValSerLysAlaLysPheAlaAlaGlySerValVal 2108
Db 1223 AATGGCATCAAGATAGGAGAC-----CGGTACATGTTGAGATGCTGCTG 1267
Qy 2109 SerPheLysCysMetGluGlyPheValLeuAsnThrSerAlaLysIleGluCysMetArg 2128

1268 TCCTTCAGTGTGAACCTGGGTATACCTTGCAGGGCGGCTCACACATTTCTTGTATGCCA 1327
129 Gly-----GlyGlnTrpAen---ProSerProMetSerIleGlnCysIleProValArg 2145
1328 GGAACGTAAGTCTGGAACATACCTTCCCTTTG-----TGCAAT---GCCACC 1375
2146 CysGlyGluProProSerIleMetAenGlyTyrAlaSerGlySerAenTyr----- 2162
1376 TGTGGTGGGACACTAACAGCATGAGTGGAGTGATCTGTAGCCAGCGTTGCCAGGCTCA 1435
2163 -----ProGluCysValProValGluCysProGlnProGlu-----GluIlePr 2447
1436 TACCCCAACAACCTTGACATGCACTCGAAGATATCACTGCCCCATTTGCTATGTCACAT 1495
2168 ValAlaTyr-----SerCysAenLysGlyPheTyr---IleLysGlyGlu 2181
1496 ATCCAAATTTCTAAATTTCTCACTGAAGCCACCATGACTACTGTGGAGATCCAGATGGC 1555
2182 LysLysSerThrCysGluAlaThrGlyGlnTrpSerPro----- 2195
1556 CCTTACCACAGCAGTCCAATGATGGGCCAGTTTCAGTGGCCCTGACCTGCTCATCATATTG 1615
2196 IleProThrCysHisProVal----- 2202
1616 CTGAGCACCACACAGAAACCTTCATCGCTTCTATAGTGACCACTCACAGAACCGGCAA 1675
2203 -----SerCysGlyGluProProLysVal 2210
1676 GGATTTAACTCGTTACCAAGCTTATGAGTTACAGAATGCCACATCCACCCCGCATTC 1735
2211 GluAenGlyPheLeuGluHisThrThrGlyArgIlePheGluSerGluValArgTyrGln 2230
1736 CAGAATGGTTCATGATCACTCGGATACAGGTGGGCCAGTGG---ATCTCTTTGAG 1792
2231 CysAenProGlyTyrLysSerValGlySerProValPheValCysGln-----AlaAen 2248
1793 TGTATATCCCGCTACATCTTGTAGGCCACCTGTCTGCTCACCCTGCTGCCAGCATGGAACCTGAC 1852
2249 ArgHisTrpHisSerGluSerProLeuMetCysValProLeuAenCysGlyLysProPro 2268
1853 AGAACTGAACATACCTTTCCACCGGTGTGACGCTCCC-----TGTGGATATAATGTG 1906
2269 ProIleGlnAenGlyPheMetLysGlyGluAenPheGluValGlySerLysValGlnPhe 2288
1907 ACATCACAGAAATGGTACCATTTATCCCTCGGTTCCTCA----- 1945
2289 PheCysAenGluGlyTyrGluLeuValGlyAenSerSerThrThrCysGlnLysSerGly 2308
1946 -----GATGAGTACCAATTTCTGAAGGACTGTCTGTGGCT----- 1980
2309 LysTrpAenLysSerAenProLysCysMetProAlaLysCysProGluProProLeu 2328
1981 -----GGTCACTGTCTCCCTCCAGGCCACCG 2004
2329 LeuGluAenGlnLeuValLeu----- 2335
2005 AGTGTACATCAATTTTCCCTGTCGACAGCCAGGCTGTAATGACTACATTTGCTGTGTG 2064
2336 -----LysGluLeuThrThrGluValGlyValThrPheSerCysLys--- 2350
2065 GGATGGTCTCTGACCAAGACTCCCTCAGCTCGGGGTCTTCAGTGGAAACACTGCCCTTGA 2124
2351 GluGlyHisValGlnGlyProSerValLeuLysCysLeuProSerGln----- 2367
2125 GACACCTACAGCTCCACCAACAGGT-----CTGTCTCAATTCACAG 2169
2368 -----GlnTrpAen-----AspSerPhePro-ValCysLysIleValLeuCy 2381
2170 CGATTTCTCCATGGAGGCTTCTTGTCTCTCAATTTCCATGTCATTTTCAGTGAAGAGGTG 2229
2381 sThrProProProLeuIleSerPheGlyValProIleProSerSerAlaLeu----- 2398

2230 CCGCCCTTCCCCCA-----GCAGTCCACAGGCTGACCTGCTTACAGA 2271
2399 -----HisPheGlySerThrValLysTyrSerCysValGlyGlyPhePheLe 2414
2272 AGATGAGGACTTTGAATAGGGGACTTCGTGAAGTACAGTGCATCCAGGTCACACATT 2331
2414 uArgGlyAenSerThrThrLeuCysGlnProAspGlyThrTrpSerSerProLeu----- 2432
2332 GTTGGGAAGCGACACCTCGACATGCAAG-----CTCAGCTCACAGCTACTGTT 2379
2433 -----ProGluCysValProValGluCysProGlnProGlu-----GluIlePr 2447
2380 CCAAGGCTTCCACCCACCTGTGAAGCACCAATGCCAGCAATGAAGTGCACAGATC 2439
2447 oAenGlyIleIleAspValGlnGlyLeuAla-----TyrLeuSerThrAlaLeuTyrTh 2465
2440 TTCTGGGTGATTTCTCAGTCCCGGTACCCAGGCACTATTTTAATCC-----CACAC 2493
2465 rCys-----LysProGlyPheGluLeuValGlyAenThrThrLe 2479
2494 ATGTGCTTGGAGTATTAAAGTGGAGCCAAACTTC-----AACATCAC-ACTCT 2540
2479 uCysGlyGluAenGlyHisTrpLeuGlyGlyLysProThrCysLysAlaIleGluCys-L 2499
2541 TTGTGGACAC-----ATTCCAAGTGAAGAAGCAATTTGATGCAC 2579
2499 euLysProLysGluIleLeuAenGlyLysPheSerTyrThrAspLeu----- 2514
2580 TG-----GAAGTGTGTGATGTTCTTCTGACAAAGTCTCTGTAGTGTCTTAA 2630
2515 -----HisTyrGlyGlnThrValThrTyrSerCysAenThrGlyPheArgLeuGlu 2532
2631 GTGGGAACACACACTGGAACAGTCCAAATTTTACCAGCAGAGTAAACCATCTGTACCTCGCG- 2689
2532 lyProSerAlaLeuThrCysLeuGluThrGlyAspTrpAspValAspAlaProSerCysA 2552
2690 -----TGTCCACACACCAACCAACAGCA 2714
2552 en-----AlaIleHisCysAspSerProGlnProIleGluA 2564
2715 AGAAAGATTCAGATTCGCTATGTCAGCTCTTACTGACGCTCACCTCTACACTCAGGA 2774
2564 sngGlyPheVal-----GluGlyAlaAspTyrSerTyrGlyAlaIleIleIle 2580
2775 ATGTGTGGCATTTTAAATAAAACCGCTGGGCA-----GTGGGAGCAAGGTACACT 2825
2580 yrSerCysPheProGlyPheGlnValAlaGlyHisAlaMetGlnThrCysGluGluSer- 2599
2826 ACTTCTGCAAGCTGTGATATCGAATGATTCGACAGCAATGCGACCTGACAGGCGGAACC 2885
2600 -----GlyTrpSerSerIleProThrCysMetProIleAspCysGlyL 2615
2886 CAGTGGCGGTGATCAGTGGGACTCGATGCGCACCACTTTGCGAGGCTGTGCTCTGTGGA 2945
2615 euProProHisIleAspPheGlyAspCysThrLysLeuLysAspGlnGlyTyrPheG 2635
2946 TTCCA----- 2950
2635 luGlnGluAspAspMetMetGluValProTyrValThrProHisProProThrHisLeuG 2655
2950 ----- 2950
2655 lYalaValAlaLysThrTrpGluAenThrLysGluSerProAlaThrHisSer----- 2672
2951 -----GAGGCTCGGGGAATGGCTCATTCACGG 2978
2673 --SerAenPheLeuTyrGlyThrMetValSerTyrThrCysAenProGlyTyrGluLeuL 2692
2979 GCAATGATGTTCACTTAGACAGTAAAGTGAATATGATGTAATGAAGGCTTCAAGCTGG 3038
2692 euGlyAen-----ProValLeuIleCysGlnGluAspGlyThrTrpAen-----GlyS 2708
3039 ATGCGAGTCAAGGAGCCACCCTGATATGTATGTAAGAAGATGGCCCTGTGGAGCAACAGAGGA 3098

Qy	3048	laAepGlyGlnTrpSerSerGlyPheProHisCysGluHisThrSerCysGlySerLeuP	3068
Db	4167	TAAATGATCTCTGTCAGGAGTACAGCCTGTGTGTAGGCGGTCTCTCGGGAACACGAC	4226
Qy	3068	roMetIleProAsnAlaPheIle-----SerGluThrSerSerTrpLysGluAsnValI	3086
Db	4227	GCACGCGCAACACGGGATGATCTCAGCAGTGTATGGATCTCTTCTCCAGCTCTGTGCA	4286
Qy	3086	leThrTyTrSerCysArgSerGlyTyTrValIleGlnGlySerSerAspLeuIleCysThrG	3106
Db	4287	TC---TATGCTGCTGGGAAGGTACAGACGCTCGGGCTCATGCTCGGCACCTGCACGG	4343
Qy	3106	luLysGlyValTrpSerGlnProTyTrProValCysGluProLeuSerCysGlySerProp	3126
Db	4344	CCAATGGGAGCTGCGACGACGACAGCTCTGACTGCACATCATCAGCTGTGGGGATCTCTG	4403
Qy	3126	roSerValAlaAsnAlaValAlaThrGlyGluAlaHisThrTyTrGluSerGluValLysL	3146
Db	4404	GCACACTGCCCAATGGCATTCAGTTTGGGACAGACTTCATCTTTCAACAGACCGTGAAGT	4463
Qy	3146	euArgCysLeuGluGlyTyTrThrMetAspThrAspThrAsp---ThrPheThrCysGlnL	3165
Db	4464	ATCAGTGCACACCCCGGTACTCTAATGGAGCCCAACATCATCCACCATCCGCTGCACCA	4523
Qy	3165	ysAspGlyArgTrpPheProGluArgIleSerCysSerProLysLysCysProLeuProG	3185
Db	4524	AAGATGCGACATGCATCAGAGCGGGCCCTCTGCAAAAGCTGTCTGTGCAACACGACGCTC	4583
Qy	3185	luAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnArgGlnValSerV	3205
Db	4584	CTCCAGTGCACCAATGGAAAGGTGAGGGGTGCAGACTTCCGATGGGGTGCAGACATAAGCT	4643
Qy	3205	alSerCysAlaGluGlyTyTrThrPheGluGlyValAsnIleSerValCysGlnLeuAspG	3225
Db	4644	ACACTGTGTGGATGGGTACTCAGCTCTCCACTCCGCGCATCTCTGTCTGTGAGGGGGTG	4703
Qy	3225	lyThrTrpGluProProPheSerAspGluSerCysSerProValSerCysGlyLysProG	3245
Db	4704	GAGTGTGG-----AAAGGAG	4718
Qy	3245	luSerProGluHisGlyPheValValGlySerLysTyTrPheGluSerThrIleIleI	3265
Db	4719	AAGTCCCTCAGTG-----CTTGGCGTCTTCA	4744
Qy	3265	yrGlnCysGluProGlyTyTrGluLeuGluGlyAsnArgGluArgValCysGlnGluAsnA	3285
Db	4745	TCCAGTGCACAGCCCCCACTGTATAGTGGATCTCTCCAGAGAACCTGCACGCCGATG	4804
Qy	3285	rgGlnTrpSerGlyGlyValAlaIleCysLysGluThr-----	3297
Db	4805	GGATATGGAGTGGCATCCAGCCCACTTGTATAGNNNNNNNNNNNNNNNNNNNNNNNNNN	4864
Qy	3297	-----	3297
Db	4865	NN	4924
Qy	3297	-----	3297
Db	4925	NN	4984
Qy	3298	-----ArgCysG	3300
Db	4985	NN	5044
Qy	3300	luThrProLeuGluPheLeuAsnGlyLysAlaAsp-----IleGluAsnArgThrT	3317
Db	5045	AGACCCCA-----GCGCATGCAGACGCTGAGAGCATCGATCTTCCAGCTT	5089
Qy	3317	brGlyProAsnValValTyTrSerCysAsnArgGlyTyTrSerLeuGluGlyProSerGluA	3337
Db	5090	TTGGCTACACTTAGTCTACACTGCCACCCCGGATTTTCTTCTGTGTGGTGGATCTGAC	5149
Qy	3337	laHis---CysThrGluAsnGlyTyTrTrpSerHisProValProLeuCysGlyLys-----	3353

Qy	768	TyrCysAlaIysThrGluAspGlyValIrrPlysProThrTyrThrGluThrProAspCys	787
Db	541	TATTGTGCTTATGAAGATGGCGTCTGGAAACCAACATATACCACTGAATGGCCAGACTGT	600
Qy	788	AlaIysLysArgPheAlaasnHisGlyPheLysSerPheGluMetPheTyrIysAlaIa	807
Db	601	GCCAAAACACGTTTTCGAAACACAGCGTTCGAAGTCTCTTGAGATGTTCTACAAACAGCT	660
Qy	808	ArgCysAspAspThrAspLeuMetLysLysPheSerGluAlaPheGluThrThrLeuGly	827
Db	661	CGTTGTGATGACACAGATCTGATGAGAGAAGTTTCTGAAGCACTTTGAGACGACCTCGGA	720
Qy	828	LysMetValProSerPheCysSerAsp	836
Db	721	AAAATGGTCCCATCATTTGTAGTGAT	747

RESULT 20				
BG534331				
LOCUS				
DEFINITION	BG534331	769 bp	mRNA	linear EST 03-APR-2001
	60253560F1	NIH_MGC_77	Homo sapiens	cDNA clone IMAGE:4663335 5',
			mRNA	sequence.

BG534331
 BG534331.1 GI:13525871
 EST.
 Homo sapiens (human)
 Homo sapiens
 Rukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 769)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.

ORIGIN

Alignment Scores:		
Pred. No.:	3,296-100	769
Score:	1382.00	251
Percent Similarity:	98.05%	1
Best Local Similarity:	97.67%	Conservative: 1
Query Match:	6.92%	Mismatches: 3
DB:	12	Indels: 2
		Gaps: 0

US-09-977-053-4 (1-3571) x BG534331 (1-769)

Qy	3163	CysGlnLysAspGlyArgTrpPheProGluAryIleSerCysSerProLysLysCysPro	3182
Db	1	TGTGAGAAAGATGGTCGGTTCCTCGACAGAAATCTCTGCGAGTCTTAAAAAATGTCTT	60
Qy	3183	LeuProGluAsnIleThrHisIleLeuValHisGlyAspAspPheSerValAsnAtrgGln	3202
Db	61	CTCCGGAAACATATAACACATATACATTGTTCATGGGAGCGATTTTCAGTGTGAATAGGCAA	120
Qy	3203	ValSerValSerCysAlaGluGlyTyThrPheGluGlyValAsnIleSerValCysGln	3222
Db	121	GTTTCTGTGTCACTGCGAGAGGGGTATACCTTTGAGGGAGTTAAACATATACGATATGTTCAG	180
Qy	3223	LeuAspGlyThrTrpGluProProPheSerAspGluSerCysSerProValSerCysGly	3242
Db	181	CTTGATGGAACTGGGAGCCACCATTCTCGATGAATCTTGCACTGCAGTTTCTTGTTGGG	240
Qy	3243	LysProGluSerProGluHisGlyPheValValGlySerIleTyThrPheGluSerThr	3262
Db	241	AAACCTGAAAGTCCAGAACATGGAATTTGTGTGTGGCAGTAAATACACCTTTGAAAGGCAC	300
Qy	3263	IleIleTyGlnCysGluProGlyTyGluLeuGluGlyValAsnArgGluAtrgValCysGln	3282
Db	301	ATTATTATCACTGTGAGCTGTGCTATGAACTAGAGGGGAAACAGGGAACTGTCTGCCAG	360
Qy	3283	GluAsnArgGlnTrpSerGlyGlyValAlaIleCysLysGluThrArgCysGluThrPro	3302
Db	361	GAGAACACAGACGTGAGTGGAGGGTGGCAATATGCAAGAGACACAGGTGTGAAACTCCA	420
Qy	3303	LeuGluPheLeuAsnGlnLysAlaAspIleGluAsnArgThrThrGlyProAsnValVal	3322
Db	421	CTTGAATTTCTCAATGGGAAAGCTGACNTTTGAAACACAGACGACTGGACCCACAGTGGTA	480
Qy	3323	TyrSerCysAsnArgGlyTySerLeuGluGlyProSerGluAlaHisCysThrGluAsn	3342
Db	481	TATTCTGCDACAGAGGTACAGTCTTTGAGGGCCATCTTGAGGCACACTGCACAGAAAT	540
Qy	3343	GlyThrTrpSerHisProValProLeuCysLysProAsnProCysProValPropheVal	3362
Db	541	GGAACTGGAGGCCACCCAGTCCCTCTCTGCACAAACCAATCCATGCCCTGTCTTTTGTG	600
Qy	3363	IleProGluAsnAlaLeuLeuSerGluLysGluPheTyThrValAspGlnAsnValSerIle	3382
Db	601	ATTCCCGAGAAAGTCTGTGCTGTGAAAGAGAGTTTTATGTGTGATCAGAAATGTGTCCATC	660
Qy	3383	LysCysArgGluGlyPheLeuLeuGlnGlyHisGly-IleIleThrCys-AsnProAspG	3402
Db	661	AAATGTAGGGAGGTTTTTCTGCTGCGAGGCCACGGCCATCATTTACTTGCACAAACCCGAGG	720
Qy	3402	luThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGly	3417
Db	721	AGAAGTGGACACAGACAGCGCCCAAAATGTGACAAATCTCATGTGT	767

RESULT	21
CF407643	
LOCUS	940 bp mRNA linear EST 02-SEP-2003
DEFINITION	CH3#049_F12T7 Canine heart normalized cDNA library in pBluescript Canis familiaris cDNA clone CH3#049 F12 5' , mRNA sequence.

CF407643
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CF407643.1 GI:34408733
EST.
Canis familiaris (dog)
Canis familiaris
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Pissipedia; Canidae; Canis.
1. (bases 1 to 940)
Yi, Y., Desai, R., Olarte, M., Henthorn, P. and George A.L.
Expressed sequence tags from Canine heart
Unpublished (2003)
Other ESTs: CH3#049_F12T3
Contact: George AL
Division of Genetic Medicine
Vanderbilt University

529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA
 Tel: 615 936 2660
 Fax: 615 936 2661
 Email: al.george@vanderbilt.edu
 Insert Length: 1564 Std Error: 0.00
 Seq primer: T7: TAATACGACTCACTATAGGG
 High quality sequence start: 38
 High quality sequence stop: 795.

FEATURES

Location/Qualifiers
 1..940
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /clone="CH3049.F12"
 /tissue_type="heart"
 /cell_type="heart"
 /dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
 /clone_lib="Canine heart normalized cDNA Library in pBluescript"
 /note="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dT primed"

ORIGIN

Alignment Scores:
 Pred. No.: 1.93e-98 Length: 940
 Score: 1362.00 Matches: 229
 Percent Similarity: 93.73% Conservative: 10
 Best Local Similarity: 89.80% Mismatches: 16
 Query Match: 6.82% Indels: 0
 DB: 14 Gaps: 0

US-09-977-053-4 (1-3571) x CF407643 (1-940)

QY 3316 ThrThrGlyProAsnValValTyrSerCysAsnArgGlyTyrSerLeuGluGlyProSer 3335
 DB 42 ACTGCTGGACTCAGGTAGTGTATCTCTGCAACAGAGGTACAGCCTTGAAGAGACCT 101
 QY 3336 GluAlaHisCysThrGluAsnGlyThrTrpSerHisProValProLeuCysLysProAsn 3355
 DB 102 GAGATACATTGCACCGACACGAGTTTGAGGCCACCCACATCTCTCTGCAACCAAT 161
 QY 3356 ProCysProValProPheValIleProGluAsnAlaLeuLeuSerGluLysGluPheTyr 3375
 DB 162 CCATGCCCTGTTCTTTTGTGATTCCTGAGAATGCTGCTCTTCTGAAAAGGAGTTTAT 221
 QY 3376 ValAspGlnAsnValSerIleLysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIle 3395
 DB 222 GTTGATCAGAACGTGTCCATCAAGTGTAGAGAAGGCTCTCTCTCCAGGGGTGGGGCATC 281
 QY 3396 IleThrCysAsnProAspGluThrTrpThrGlnThrSerAlaLysCysGluLysIleSer 3415
 DB 282 ATTACTGTAAACCCGATGAGACATGAGCGGTGACAAATGCCAAATGTGAAAATATCC 341
 QY 3416 CysGlyProProAlaHisValGluAsnAlaIleAlaArgGlyValHisTyrGlnTyrGly 3435
 DB 342 TGTGGTCCACCACTCATGTAGAAATGCGATGCTCGAGGCATACATTATCAGTATGG 401
 QY 3436 AspMetIleThrTyrSerCysTrpSerGlyTyrMetLeuGluGlyPheLeuArgSerVal 3455
 DB 402 GACATGATCACTACTCATGTGTACAGTGGATCATGTGTGGAGGGCTCCCTAGAGGATAT 461
 QY 3456 CysLeuGluAsnGlyThrTrpThrSerProIleCysArgAlaValCysArgPhePro 3475
 DB 462 TGCCTAGAGATGGAACTGGACATCACTCTATTGTCAGAGCTGTCTGTGATCCCA 521
 QY 3476 CysGlnAsnGlyGlyIleCysGlnArgProAsnAlaCysSerCysProGluGlyTrpMet 3495
 DB 522 TGTCAAGATGGAGGTATCTGCCAACGCCCAATGCTGTGTTCTCTGTCAGATGCTGGATG 581

QY 3496 GlyArgLeuCysGluGluProIleCysIleLeuProCysLeuAsnGlyGlyArgCysVal 3515
 DB 582 GGGGGCTCTGTGAAGAGCAATATGCAATCTCTCCCTGTTTGAATGTGTGCTGTGTA 641
 QY 3516 AlaProTyrGlnCysAspCysProProGlyTyrThrGlySerArgCysHisThrAlaVal 3535
 DB 642 GCCCCTTACCACTGTGACTGTCCACCTGGCTGGACGGGTCCCGCTGTCTATACAGCTGT 701
 QY 3536 CysGlnSerProCysLeuAsnGlyGlyLysCysValArgProAsnArgCysHisCysLeu 3555
 DB 702 TGCCAGTCTCCCTGCTTAAATGTGGAAATGTGTAGACCAACCATGTCTATGTGTCT 761
 QY 3556 SerSerTrpThrGlyHisAsnCysSerArgLysArgArgThrGly 3570
 DB 762 TCAGCTTGGACAGGACATGACTGCTCAGNGAAGAGGAGACTGG 806

RESULT 22

AL706149 738 bp mRNA linear EST 04-SEP-2003
 LOCUS DKFZp686E243_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 DEFINITION DKFZp686E243_5', mRNA sequence.

ACCESSION AL706149

VERSION AL706149.1 GI:19689504

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 738)

AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and

Wiemann,S.

TITLE EST (Duesterhoeft, et al.)

JOURNAL Unpublished (1999)

COMMENT Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing

consortium of the German Genome Project.

No a1 sequence available.

This clone (DKFZp686E243) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1..738

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp686E243"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="686 (synonym: hlcc3)"

/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;

cDNA-collection"

ORIGIN

Alignment Scores:
 Pred. No.: 2.16e-98 Length: 738
 Score: 1359.00 Matches: 238
 Percent Similarity: 98.36% Conservative: 2
 Best Local Similarity: 97.54% Mismatches: 4
 Query Match: 6.80% Indels: 0
 DB: 9 Gaps: 0

US-09-977-053-4 (1-3571) x AL706149 (1-738)

QY 1253 PheIleCysGluCysProSerGlyTyrThrGlyLysArgCysGluAsnIleAsnGlu 1272
 DB 3 TTTCATTGTGTGAGTTTCCATCAGGTTCACAGGTCAGCGGTGTGAGAAAAATATAAATGAG 62

QY 1273 CysSerSerProCysLeuAenGlyIleCysValAspGlyValAlaGlyTyrArg 1292
 Db 63 TGTAGCTCAGTCCCTGTTTAAATAAGGAATCTGTGTTGATGGTGGCTGCTATCGT 122

QY 1293 CysThrCysValGlyPheValGlyLeuHisCysGluThrGluValAenGluCysGln 1312
 Db 123 TGCACATGTGTGAAGGATTTGTAGCCCTGCATTGTGAGCAGCAANTCNTGAATGCCAG 182

QY 1313 SerAenProCysLeuAenAenAlaValCysGluAspGlnValGlyGlyPheLeuCysLys 1332
 Db 183 TCAAAACCATGCTTAAATAATGAGTCTGTGAAGACCAAGTGTGGGGGATCTTGTGCAAA 242

QY 1333 CysProProGlyPheLeuGlyThrArgCysGlyLysAenValAspGluCysLeuSerGln 1352
 Db 243 TGCCCACTGGATTTTGGGTACCCGATGTGAAGAAGCTGATGAGTGTCTCAGTCAG 302

QY 1353 ProCysLysAenGlyAlaThrCysLysAspGlyAlaAenSerPheArgCysLeuCysAla 1372
 Db 303 CCATGCAAAATGGAGCTACCTGTAAAGACGGTGGCCATAGCTTCAGATGCTGTGTGCA 362

QY 1373 AlaGlyPheThrGlySerHisCysGluLeuAenLysAenGluCysGlnSerAenProCys 1392
 Db 363 GCTGGCTTCACAGGATCACACTGTAATTGAACATCAATGAATGTCAGTCTAATCCATGT 422

QY 1393 ArgAenGlnAlaThrCysValAspGluLeuAenSerTyrSerCysLysCysGlnProGly 1412
 Db 423 AGAAATCAGGCACATGTGTGATGAATTAATTCATACATGTTGTAATATGTGAGCAGGA 482

QY 1413 PheSerGlyGlnArgCysGluThrGluGlnSerThrGlyPheAenLeuAspPheGluVal 1432
 Db 483 TTTTCAGGCAGAAAGGTGTGAACAGACAGTCTACAGGCTTTAACTGGATTTTGAAGTT 542

QY 1433 SerGlyIleTyrGlyTyrValMetLeuAspGlyMetLeuProSerLeuHisAlaLeuThr 1452
 Db 543 TCTGTATCTATGATATGTCTATGCTAGATGGCTGCTGCCATCTCTCCATGCTCTAACC 602

QY 1453 CysThrPheTrpMetLysSerSerAspAspMetAenTyrGlyThrProLysSerTyrAla 1472
 Db 603 TGTACCTTCGTGATGAATCCTCTGACGACATGAATATGGAACACCAATCTCTCTATGCA 662

QY 1473 ValAspAenGlySerAspAenThrLeuLeuLeuThrAspTyrAsnGlyTyrValLeuTyr 1492
 Db 663 GTTGATACGGCAGCAGACATACCTTGCTCTGCTGACTGATTATTAACGGCTGGTCTTTAT 722

QY 1493 ValAenGlyArg 1496
 Db 723 GTGAATGGCAGG 734

RESULT 23

CP409971 839 bp mRNA linear EST 02-SEP-2003
 LOCUS CH3#063_H05MF Canine heart normalized cDNA Library in pBluescript
 DEFINITION Canis familiaris cDNA clone CH3#063_H05 5', mRNA sequence.

ACCESSION CP409971

VERSION CP409971.1

KEYWORDS EST; GI:34411217

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

1 (bases 1 to 839)

Yi.Y., Desai.R., Olarte.M., Henthorn.P. and George A.L.

Expressed sequence tags from Canine heart

Unpublished (2003)

Other ESTs: CH3#063_H05MR

Contact: George AL

Division of Genetic Medicine

Vanderbilt University

529 Light Hall, 2215 Garland Avenue, Nashville, TN 37232-0275, USA

Tel: 615 936 2660

Fax: 615 936 2661

Email: al.george@vanderbilt.edu

Insert Length: 1614 Std Error: 0.00

Seq primer: MF: GTTTTCCAGTCACGACGTTG
 High quality sequence start: 83
 High quality sequence stop: 753.

FEATURES

source

1..839
 /organism="Canis familiaris"
 /mol_type="mRNA"
 /db_xref="taxon:9615"
 /clone="CH3#063_H05"
 /tissue_type="heart"
 /cell_type="heart"
 /dev_stage="mixed developmental stages (adult, 30 day - 40 day fetal)"
 /clone_lib="Canine heart normalized cDNA Library in pBluescript"
 /note="Organ: heart; Vector: pBluescript; Site 1: 5' of vector NotI; Site 2: 3' of vector EcoRI; Tissue source: dog heart (adult, 30 day - 40 day fetal), right and left atria and ventricle. Dog breed - mixed (beagle, German shepherd, pointer, Irish setter). Library construction: oligo-dr primed"

ORIGIN

Alignment Scores:

Pred. No.: 4,51e-95 Length: 839
 Score: 1319.00 Matches: 229
 Percent Similarity: 90.19% Conservative: 10
 Best Local Similarity: 86.42% Mismatches: 24
 Query Match: 6.60% Indels: 2
 DB: 14 Gaps: 0

US-09-977-053-4 (1-3571) x CP409971 (1-839)

QY 3303 LeuGluPheLeuAenGlyLysAlaAspIleGluAenArgThrThrGlyProAenValVal 3322
 Db 48 TTGGAGCTCCACCGCGGTGGCGCGCGCTCGAGCAGCAGCTGCTGGACTCAGGAGTG 107

QY 3323 TyrSerCysAenArgGlyTyrSerLeuGluGlyProSerGluAlaHisCysThrGluAen 3342
 Db 108 TATTCCTGCACAGAGGGTACAGCCTTGAAGAGGACACCTGAGATACATTCGCCGACAC 167

QY 3343 GlyThrTrpSerHisProValProLeuCysLysProAenProCysProValProPheVal 3362
 Db 168 GGAGTTTGGAGCCACCATCTCTCTGCANACCAATCCATGCTGCTCTTTTGTG 227

QY 3363 IleProGluAenAlaLeuLeuSerGluLysGluPheTyrValAspGlnAenValSerIle 3382
 Db 228 ATTCTCTGAGAAATGCTGCTCTTCTGAAAAGGAGTTTATTTGATCAGAACGTCCTATC 287

QY 3383 LysCysArgGluGlyPheLeuLeuGlnGlyHisGlyIleIleThrCysAenProAspGlu 3402
 Db 288 AAGTGTAGAGAGGCTTCTCTCCAGGTTGGGGGATCATTTACCTGTAAACCCGATGAG 347

QY 3403 ThrTrpThrGlnThrSerAlaLysCysGluLysIleSerCysGlyProProAlaHisVal 3422
 Db 348 ACATGACGGTGCACAAATGCCAAATGTGAATAATATCTGTGTCTCCACCACTCATGTA 407

QY 3423 GluAenAlaIleAlaArgGlyValHisTyrGlnTyrGlyAspMetIleThrTyrSerCys 3442
 Db 408 GGAATGCGATGCTGTCGAGGGCATACATTTATCATGATGAGGACATGATCATCATCATGT 467

QY 3443 TyrSerGlyTyrMetLeuGluGlyPheLeuArgSerValCysLeuGluAenGlyThrTrp 3462
 Db 468 TACAGTGGATACATGTTGGAGGGCTCCCTTAAGAGAGTATTTGCTAGAGATGGACCTGG 527

QY 3463 ThrSerProProIleCysArgAlaValCysArgPheProCysGlnAenGlyGlyIleCys 3482
 Db 528 ACATCACCTCTCTATTGTCAGAGCTGCTGTGTGATTCCTCATGTCAGATGGATGATCTGC 587

QY 3483 GlnArgProAenAlaCysSerCysProGluGlyTrpMetGlyArgLeuCysGluGluPro 3502
 Db 588 CAACGCCCAATGCTTGTCTCTGTCAGATGGCTGGATGGGGCGGCTCTGTGGAAGGCCA 647


```

http://image.llnl.gov
Plate: LLCM1591 row: 1 column: 06
High quality sequence stop: 738.
Location/Qualifiers
1..756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4732181"
/lab host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 79"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcgctcggcc); Site_2: SfiI
(ggcgcattggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

```

ORIGIN

```

Alignment Scores:
Pred. No.:          3.18e-93
Score:              1295.00
Length:             756
Matches:            250
Percent Similarity: 98.04%
Conservative:       0
Best Local Similarity: 98.04%
Mismatch:           1
Query Match:        6.48%
Indels:             5
DB:                 12
Gaps:               0

```

US-09-977-053-4 (1-3571) x BG620159 (1-756)

476	AspLysLeuThrCysGlnGlyAsnSerGlnTrpAspGlyProGluProArgCysValGlu	495
1	GATAAGCTTACTTGTCAAGGAAACAGCCAGTGGGATGGCCAGAACCCCGGTGTCTGGAG	60
496	ArgHisCysSerThrPheGlnMetProLysAspValIleIleSerProHisasnCysGly	515
61	CGGCACCTGTTCACCTTTCAGATGCCAAGAATGTCATCATATCCCCCACAACTGTGGC	120
516	LysGlnProAlaLysPheGlyThrIleCysTrpValSerCysArgGlnGlyPheIleLeu	535
121	AAGCAGCCAGCCAAATTTGGGACGATCTGCTATGTAAAGTTCGCCCAAGGGTTCATTTTA	180
536	SerGlyValLysGluMetLeuArgCysThrThrSerGlyLysTrpAsnValGlyValGln	555
181	TCTGGAGTCAAGAGAAATGCTCAGATGTACCACTTCTCGAAATATGGAATGTCTGGAGTTCAG	240
556	AlaAlaValCysLysAspValGluAlaProGlnIleasnCysProLysAspIleGluAla	575
241	GCAGCTGTGTGAAGACGTGGAGGCTCCTCAATCAACTGTCTTAAAGACATAGAGGCT	300
576	LysThrLeuGluGlnGlnAspSerAlaAsnValThrTrpGlnIleProThrAlaLysAsp	595
301	AAGACTCTGGACAGCAAGATTCTGCCAATGTACTTGGCAGATTTCCAAACAGCTAAAGAC	360
596	AsnSerGlyGluLysValSerValHisValHisProAlaPheThrProProTrpLeuPhe	615
361	AACTCTGGTGAAGAGGTCTCAGTCCACCGTTCAATCAGCTTTTCAACCCACCTTACCTTTTC	420
616	ProIleGlyAspValAlaIleValTyrrAlaThrAspLeuSerGlyAsnGlnAlaSer	635
421	CCAATTGGAGATGTGTCTATCGTATACACGGCAACTGACCTTATCCGGCAACAGGCCACG	480
636	CysIlePheHisIleLysValIleAspAlaGluProProValIleAspTrpCysArgSer	655
481	TGCATTTTCCATATCAAGGTTATTGATGCAGAACCCACTGTCATAGACTGGTGCAGATCT	540
656	ProProValGlnValSerGluLysValHisAlaAlaSerTrpAspGluProGlnPhe	675
541	CCACCTCCCGTTCAGGCTCTCGGAGAGGTATCATGCCGAAGTGGCAATGAGCTCAGTTC	600